REVIEW

of the

TRICAL COMMUNICATION LABORATORY

TELEGRAPH AND TELEPHONE PUBLIC CORPORATION

Former, Reports of the Electrical Communication Laboratory
Nippon Telegraph and Telephone Public Corporation
from Vol.1, No.1, Sept. 1953 to Vol.7, No.12, Dec. 1959

Published Bimonthly by

ELECTRICAL COMMUNICATION LABORATORY
PON TELEGRAPH AND TELEPHONE PUBLIC CORPORATION
1551, Kitizyôzi, Musasino-si, Tôkyô, Japan

VOLUME 9
NUMBERS 5-6

May - June
1961

Review of the Electrical Communication Laboratory

EDITORIAL COMMITTEE

Susumu OKAMURA *Chairman*

Tadasu FUKAMI
Giichi ITO
Hideo KAWASAKI
Yukio NAKAMURA
Tatsuya SHIMBORI
Takuzo SHINDO
Toyotaro SHIRAMATSU
Takakichi UZAWA
Ginsaku YASAKI

Nobukazu NIIZEKI

EDITORIAL STAFF

Takaaki MIWA *Editor*
Tutomu SAITO *Assistant Editor*
Masao HIROSE 〃

The Review of the Electrical Communication Laboratory is published six times a year (bimonthly) by the Electrical Communication Laboratory, Nippon Telegraph and Telephone Public Corporation, 1551, Kitizyôzi, Musasino-si, Tôkyô, Japan. Telephone: Tôkyô (391) 2261 and 2271.

All rights of republication, including translation into foreign languages, are reserved by the Electrical Communication Laboratory, NTT.

Subscriptions are accepted at ¥1,200 per year. Single copy ¥300 each. Foreign postage is ¥400 per year or ¥65 per copy. Remittance should be made in check payable to the Electrical Communication Laboratory and sent to the Director.

Printed by Aoyama Printing Co., Tokyo, Japan

R E V I E W
of the
ELECTRICAL COMMUNICATION LABORATORY

NIPPON TELEGRAPH AND TELEPHONE PUBLIC CORPORATION

Volume 9, Numbers 5-6, May-June, 1961

U.D.C. 519.2

Majority Decision Functions of Six Variables*

Iwao TODA,† **Mamoru KONDO,**† and **Saburo MUROGA**‡

A method to obtain all majority decision functions is described. A majority decision function is a logical function which can be represented by a single majority decision element, such as a parametron or a magnetic core.

It is shown that given a majority decision function, its canonical representative can easily be determined from the intrinsic properties of majority decision functions.

Thus we can obtain a set of possible condidates for the canonical representatives which contains all of the true representatives. Then we can pick up all the true representatives from these candidates by a method of linear programming.

The method proposed is proved to be quite efficient as compared with the conventional method.

By this method a table of all the majority decision functions of six variables is obtained together with their optimum structures. Also, several properties of majority functions which have been induced from the table are given.

1. Introduction

Recently, in the field of electronic computers there have been widely adopted majority decision logical elements such as parametrons and magnetic cores, whose operation is based upon the principle of majority decision among input quantities.

It is easily shown that with the concept(1) of coupling weights of input variables introduced by one of the authors the majority decision element can be generalized to include so-called threshold logical elements.(2) Hence majority decision elements are now considered to be important elements for the realization of logical networks.

One of the prominent features of majority decision elements is that a single majority decision element can represent a wide variety of complex logical functions with a slight change of its structure.(1)(2)(3) Recently a hypothesis was proposed that the great discrepancy between the memory capacity of a human brain and the number of neurons in it is attributed to this property of majority decision elements.

Considering this importance of majority decision elements, we have reported a series

---

* MS in Japanese received by the Electrical Communication Laboratory, on 26 August, 1960. Originally published in the *Kenkyû Zituyôka Hôkoku* (*Electrical Communication Laboratory Technical Journal*), N.T.T., Vol. 10, No. 3, pp. 369-403, 1961.

† Electronic Research Section.
‡ IBM Research Center (Mohansic Lab.) New York.

of studied on the majority decision function, which is a logical function represented by a single majority decision element.(1)(2)(3)(5)

There rather restricting necessary conditions for majority decision functions(2) have been obtained. Also a way to decide by the method of linear programming whether a given logical function is a majority decision function and at the same time obtain an optimum structure of the majority decision element realizing the given function has been proposed.(5)

From the above we can gain a lucid insight into the properties of the majority decision function, but the criterion of the majority decision function by linear programming is too mechanical to believe we have established sufficient conditions for the majority decision function satisfactorily.

With this sufficient condition as the nucleus there exist several problems to be solved. In this stage of research it is concluded that it is profitable to tabulate all the majority decision functions of several variables from both the practical and the theoretical points of view. This is a main motive of this study.

The table of all majority decision functions of the variables equal to or less than five has already obtained by one of the authors.(3) Therefore all majority decision functions of six variables were sought in this study.

S. Muroga(3) obtained all of the majority decision functions of five or less variables by picking up the distinct functions among all the functions generated under the condition that the total number of coupling weights does not exceed some prescribed number. This method, easily programmed on an electronic computer, has the defects that it cannot be guaranteed to be able to obtain all the majority decision functions, and that the same function may be generated a number of times slowing down the computing speed. To apply this method for the case of six variables would require much time and is almost impracticable.

To eliminate this difficulty a new method is proposed in this study. This method of obtaining all majority decision functions of η variables consists of two stages. In the first stage a set of possible candidates for the canonical representatives, which contains all of true representatives, are obtained by making use of the table of the majority decision functions of n-1 variables and of the properties of majority decision functions of n variables. In the second stage each possible candidate is tested to see if it is a true candidate by the method of linear programming. This method, which never misses any canonical representative, is quite efficient; for example, all the candidates of six variables were proved to be true representatives, and it is easily programmed on an electronics computer. Actually we obtained the table of majority decision functions of six variables with the aid of the parametron computer, MUSASINO-1, in our laboratory.

Canonical representatives of majority decision functions, a procedure to obtain all the majority decision functions, and a table of majority decision functions of six variables with their optimum structures are discussed in the following sections.

2. Majority Decision Element and Majority Decision Function

A majority decision element of n variables is a logical element with n Boolean inputs, x_1, $x_2 \cdots\cdots x_n$, and one Boolean output, which is an output one for those combinations of inputs satisfying (1.a) below and an output zero for those combinations of input satisfying (1.b):

$$\sum_{i=1}^{n} w_i X_i \geq T \qquad (1.\mathrm{a})$$

$$\sum_{i=1}^{n} w_i X_i < T \qquad (1.\mathrm{b})$$

where w_i is a prescribed constant called the "coupling weight" associated with the input X_i and T is also a prescribed constant called the "threshold."(1)(2)(5)

For parametrons or magnetic cores the coupling weight w_i corresponds to the number of turns of the winding for the input x_i. The threshold T is related to w_c of the constant input by the relation shown below:

$$w_c = \sum_{i=1}^{n} w_i + 1 - 2T \qquad (2)$$

where $W_c \geq 0$ means the constant of one is added to the element and $W_c \leq 0$ means the constant is zero.

As a majority decision element can be completely determined by $n+1$ real numbers $w_1, w_2, \cdots\cdots, w_n : T$, a set of $n+1$ real numbers $[w_1, w_2, \cdots\cdots, w_n : T]$ will be called a structure of the majority decision element.

A logical function which a single majority decision element represents is called a majority decision function.

For example, a majority decision element with the structure $[2, 1, 1 : 2]$ denotes the logical function $x_1x_2x_3$, hence this function is a majority decision function by definition. On the other hand, $x_1x_2+x_3x_4$ is not a majority decision function, for it can never be realized by any single majority decision element.

3. Properties of Majority Decision Functions

In this section various properties of majority decision functions which will be utilized in this paper are quoted without proof. The proof will be found in ref. (2) and (3). Throughout this section the structure of a majority decision function $f(r_1, \cdots\cdots, x_n)$ will be denoted by $[w_1, w_2, \cdots\cdots, w_n : T]$.

1) The function $f(x_1, \cdots\cdots, x_n)$ derived from a majority decision function $f(x_i, \cdots\cdots, x_n)$ by negating x_1 is also a majority decision function and can be realized by the majority decision element $[-w_1, w_2, \cdots\cdots, w_n : T-w_1]$

2) If $w_1 \geq 0, w_2 \geq 0, \cdots\cdots, w_n \geq 0$; then $f(x_1, \cdots \cdots x_n)$ can be expressed as a polynomial of $x_1, x_2, \cdots\cdots, x_n$ without any negated variables. Such a polynomial will be called a positive polynomial in $x_1, x_2, \cdots\cdots, x_n$.

From 1) and 2), it can be seen that we are concerned solely with those majority decision functions such that $w_i \geq 0$ for each i, or with those majority decision functions of positive polynomials in $x_1, x_2, \cdots\cdots, x_n$. Only such majority decision functions will be considered throughout this section.

3) A positive majority decision function can be expressed as

$$f(x_1, \cdots\cdots, x_n) = M(x_2, \cdots\cdots, x_n)x_1 + N(x_2, \cdots\cdots, x_n) \qquad (3)$$

where M and N are both positive majority decision functions in $x_2, x_3, \cdots\cdots x_n$.

4) If $w_1 \geq w_2, w_3, \cdots\cdots w_n$; then

$$M(x_2, \cdots\cdots, x_n) \supseteq N(x_2, \cdots\cdots, x_n) \qquad (4)$$

holds, where $\supseteq$ denotes the implication relation of Boolean functions.⁽²⁾

5) Define the binary relation $x_i \gtrsim x_j$ for any positive logical function $g(x_1, \cdots\cdots, x)$ to mean that

$$g(x_1, \cdots, \overset{\overset{i}{\vee}}{1}, \cdots, \overset{\overset{j}{\vee}}{0}, \cdots x_n) \supseteq g(x_1, \cdots, \overset{\overset{i}{\vee}}{0}, \cdots, \overset{\overset{j}{\vee}}{1}, \cdots, x_n) \qquad (5)$$

holds. Then this relation can be shown to be a partial ordering; namely, if $x_1 \gtrsim x_2$ and $x_2 \gtrsim x_3$, then $x_1 \gtrsim x_3$ follows.

For completeness,

$$x_i \sim x_j \qquad (6)$$

will be defined.

6) For all the input variables of a majority decision function the relation defined in 5) is a total ordering and is homomorphic to the ordinal relation of w's as real numbers. That is:

$$\text{if} \quad w_1 > w_2, \quad x_1 \gtrsim x_2; \qquad (7)$$

$$\text{and if} \quad w_1 = w_2, \quad x_1 \sim x_2. \qquad (8)$$

7) If $x_1 \gtrsim x_2 \gtrsim, \cdots\cdots, \gtrsim x_n$ for majority a decision function $f(x_1, \cdots\cdots, x_n)$; namely, $w_1 \geq w_2 \geq \cdots\cdots, w_n$, then:

$$f(x_1, \cdots\cdots, x_n) = M(x_2, \cdots\cdots, x_n)x_1 + N(x_2, \cdots\cdots, x_n) \qquad (9)$$

$$\left.\begin{aligned} M(x_2, \cdots\cdots, x_n) &= m_1(x_3, \cdots\cdots, x_n)x_2 \\ &\quad + m_2(x^3, \cdots\cdots, x_n) \\ N(x_1, \cdots\cdots, x_n) &= n_1(x_3, \cdots\cdots, x_n)x_2 \\ &\quad + n_2(x_3, \cdots\cdots, x_n) \end{aligned}\right\} \qquad (10)$$

where m_1, m_2, n_1, n_2 are positive majority decision functions in $x_3, x_4, \cdots\cdots, x_n$ for which

$$m_1 \supset m_2 \supseteq n_1 \supset n_2 \qquad (11)$$

holds.

8) Denoting the dual function of $f(x_1, \cdots\cdots x_n)$ by $f^*(x_1, \cdots\cdots, x_n)$; if $f(x_1, \cdots\cdots, x_n)$ is a majority decision function then $f^*(x_1, \cdots\cdots, x_n)$ is also a majority decision function with the structure of $\left[w_1, \cdots\cdots, w_n : \sum_{i=1}^{n} w_i - T + 1 \right]$.

9) For a majority decision function either

$$f \supseteq f^* \qquad \text{or} \qquad f^* \supseteq f \qquad (12)$$

is true. In the former case w_c defined in (2) is positive and in the latter case w_c is negative.

4. Criterion of a Majority Decision Function and Determination of its Structure

In this section a method of determining whether a given logical function is a majority decision function or not is discussed. A detailed explanation of this section is found in ref. (5).

A logical function $g(x_1, \cdots\cdots, x_n)$ is a mapping which gives one or zero for all 2^n combinations of input variables $x_1, \cdots\cdots, x_n$.

If it is a majority decision function, there exists a set of $n+1$ real numbers $w_1, w_2, \cdots\cdots w_n$, T for which

$$\sum_{i=1}^{n} w_i X_i \geq T \qquad (13.\text{a})$$

is true for those combinations of inputs that given an output one and

$$\sum_{i=1}^{n} w_i X_i \leq T-1 \qquad (13.\text{b})$$

is true for those that give an output zero. The term (-1) on the right-hand side of equation (13.b) is a normalizing factor, inserted to eliminate the ambiguity of constant multipliers for the w's and T, which follows from definition (1).

For any one combination of inputs, either (13.a) or (13.b) holds, so that 2^n inequalities for the w's and T are obtained. An example of this is shown as follows:

$$\left\{\begin{array}{l} w_1 + w_2 + \cdots\cdots \quad + w_n \geq T \\ w_1 \qquad + w_3 \cdots\cdots + w_n \geq T \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ w_1 + w_2 \qquad\qquad\quad \leq T-1 \\ \qquad\qquad\qquad 0 \leq T-1. \end{array}\right. \qquad (14)$$

It is clear that if the system of inequalities has a finite solution, $g(x_1, \cdots\cdots, x_n)$ is a majority decision function and that the solution gives a structure to realize $g(x_1, \cdots\cdots, x_n)$.

Given a system of inequalities, the existence of its solution can be investigated without serious difficulty. But the problem of predicting its existence directly from the logical function $g(x_1, \cdots\cdots, x_n)$ has not yet been solved.

Among 2^n inequalities redundant conditions can be removed by the following procedure.

Procedure 4.1.

Let $X_1 = (x_1^{(1)}, x_2^{(1)}, \cdots\cdots, x_n^{(1)})$ and $X_2 = (x_1^{(2)}, x_2^{(2)}, \cdots\cdots, x_n^{(2)})$ be two combinations of inputs for which $g(x_1, \cdots\cdots, x_n)$ gives an output one. Each component of X_1 and X_2 is either one or zero. Suppose $x_i^{(1)} = x_i^{(2)}$ for $i = i_1, i_2, \cdots\cdots, i_m$, and $x_i^{(1)} \neq x_i^{(2)}$ for $i \neq i_1, i_2, \cdots\cdots, i_m$. From the function $g_1(x_{i_1}, \cdots\cdots, x_{i_m})$ by substituting $x_i = x_i^{(1)}$ for $i \neq i_1, \cdots\cdots, i_m$ into $g(x_1, \cdots\cdots, x_n)$. Similarly from the function $g_2(x_{i_1}, \cdots\cdots, x_{i_m})$ by substituting $x_i = x_i^{(2)}$ for $i \neq i_1, i_2, \cdots\cdots, i_m$.

Remove the inequality in (14) corresponding to X_1 if $g_1 \supset g_2$ holds.

Next let X_1 and X_2 be two combinations of inputs which make $g(x_1, \cdots\cdots, x_n)$ zero. Define g_1 and g_2 in the same manner described above. Then remove the inequality corresponding to X_2 if $g_1 \subset g_2$ holds.

Repeat the above procedure for all possible pairs of X_1 and X_2.

By this procedure we can reduce the number of inequalities to some extent. The system of inequalities reduced in this manner will be called the reduced system of inequalities of $g(x_1, \cdots\cdots, x_n)$.

For example, let $g(x_1, \cdots\cdots, x_4) = x_1x_2 + x_1x_3 + x_1x_4 + x_2x_3$, $X_1 = (1.1.1.1)$, $X_2 = (1.1.0.0)$, $X_3 = (0.1.1.0)$, each of which makes $g(x_1, \cdots\cdots, x_4)$ one. In comparing X_1 with X_2, $g_1 = 1$ and $g_2 = x_1x_2$; therefore the inequality of X_1 can be dispensed with. Comparing X_2 with X_3, we obtain $g_2 = x_2 + x_4$, $g_3 = x_2$ and $g_2 \supset g_3$, so that the inequality for X_2 can be removed. Repeating this procedure the initial $2^4 = 46$ inequalities can at last be reduced to five inequalities in this example.

In realizing a majority decision element, it is quite desirable to find an optimum structure for the given function in some prescribed sense, for example with a minimum number of windings in the case of parametrons and magnetic cores.

When a given majority decision function is positive, the total number of windings including the windings for the constant input is represented as:

$$K = \sum_{i=1}^{n} w_i + |w_c| = \begin{cases} 2T-1 & (w_c \leq 0) \\ 2\left(\sum_{i=1}^{n} w_i - T\right) + 1 & (w_c \geq 0) \end{cases} \tag{15}$$

Then the problem to be solved can be stated as follows.

Problem 4.1.

Find a set of $w_1, \cdots\cdots, w_n, T$ that minimizes the number of the total coupling weight K under the condition that they must satisfy the reduced system of inequalities for the given function.

This is apparently a typical example of linear programming, and can be solved effectively.

If the given function is not a majority decision function, the systemof inequalities has no solution and therefore Problem 4.1. has no solution.

Thus we can decide whether a given function is a majority decision function according to the existence of the solution of Problem 4.1.

When a solution does exist, the solution has the structure of the majority decision element that realizes the given function most economically.

Practically, the ambiguity of the defining formula for K in (15) is out of question, for from property (8) in the preceding section we can tell beforehand the polarity of w_c using the inclusion relation between g and g^*.

5. Classification of Majority Decision Functions

The functions which can be derived from a given logical function by the operations given below are defined as equivalent functions of the original function.

Operation 5.1.

1) Negation of input variables
2) Permutation among input variables
3) Negation of output.

It is well known that logical functions can be classified into equivalent classes by this equivalent relation.

It is readily proved that the function derived from a majority decision function by Operation 5.1. is also a majority decision function and that no majority decision function can be derived from any non-majority-decision function by the above operations. Therefore, the whole of the majority decision functions is an aggregate of several of the equivalent classes just defined.

When the structure of one function in a certain equivalent class is known, the structures of the rest can easily be obtained.

The number of equivalent classes of logical functions of n variables becomes exceedingly small compared with the total number of the functions as n increases.(6)

This holds true for majority decision functions and it is much less laborious to determine the structures of the representatives of the equivalent classes than to determine those of the whole of the majority decision functions.

In the case of general logical functions, however, it is a difficult problem to extract one representative from an equivalent class but in the case of majority decision functions it is quite easy to define a canonical representative of an equivalent class from the intrinsic nature of majority decision functions. Moreover, when a majority decision function is given the canonical representative of the class to which the given function belongs can easily be determined.

A method of determining the canonical representative of an equivalent class of majority decision functions is shown below.

From 1) and 2) in section 3 any majority decision function can be expressed as a positive polynomial by suitable negation of input variables. If none of the variables is vacuous, the way of negation to obtain the positive polynomial from the given function is unique and hence by requiring the positive polynomial for the representative the ambiguity due to 1) in Operation 5.1. can be removed. As is easily seen, this condition is equivalent to the condition that the canonical representative is always realized by a majority decision element with positive coupling weights.

The ambiguity due to 2) in Operstion 5.1. can be removed from properties 5) and 6) in section 3. All the input variables of a majority decision function can be ordered by the relation $\gtrsim$. By suitable permutation amang input variables it is always possible to rearrange the variables so that $x_1 \gtrsim x_2 \cdots\cdots \gtrsim x_n$ holds. When no vacuous variable appears, this permutation can be uniquely determined except the permutations among; for example, $x_1, \cdots\cdots, x_m$, for which $x_1 \sim x_2 \cdots\cdots \sim x_m$ holds. But $x_1 \sim x_2 \cdots\cdots \sim x_m$ means that the given function is symmetric with respect to the variables $x_1, \cdots\cdots, x_m$, hence the function is invariant under any permutation among $x_1, \cdots\cdots, x_m$. Therefore, the function for which $x_1 \gtrsim x_2 \cdots\cdots \gtrsim x_n$ holds is unique and can well be adopted as a representative. Of course this means that $W_1 \geq W_2 \cdots\cdots \geq W_n$ holds for the representative element.

There exist now two functions in each equivalent class which satisfy the two conditions just imposed. We denote one of them by f the other is f^*; that is, the dual function of f. Either $f^* \sqsupseteq f$ or $f \sqsupseteq f^*$ holds between f and f^* for majority decision functions from 9) in section 3. Therefore if we choose either of the two conditions, the representative of the equivalent class is now uniquely fixed. Let the former condition be adopted, so as to simplify the target function K in Problem 4.1. The condition $f^* \sqsupseteq f$ means $W_c \leq 0$ from 9) in section 3.

Now it is shown that there exists a unique canonical representative in each equivalent class of majority decision functions which satisfies the following three conditions.

Condition 5.1.
1) A positive polynomial
2) $x_1 \gtrsim x_2 \cdots\cdots \gtrsim x_n$
3) $f^* \sqsupseteq f$.

6. A Method to obtain All Maiority Decision Functions

With the preceding three sections as preliminaries a method is proposed here to obtain all majority decision functions.

As described in section 4, it can be determined by linear programming whether a given function is a majority decision function; and therefore, it is possible at least in principle to obtain all majority decision functions by applying the criterion to all of 2^{2^n} logical functions of n variables. But it will require an impracticably large amount of time to solve 2^{2^n} of linear programming for n of large values.

Accordingly, a method is proposed to reduce the number of functions to be tested by deleting those apparently non-majority-decision-functions from the properties of majority decision functions and by picking up one candidate for each equivalent class.

The conditions to delete the non-majority-decision functions must be simple enough to be programmed on an electronic computer.

From this point of view, the first condition chosen is that condition 5.1. must be defined for each candidate. This is quite natural from the policy that only canonical representatives are tested by linear programming.

A positive majority decision function of n variables can be represented as Mx_1+N where M and N are appropriate majority decision functions of $(n-1)$ variables.

The second condition is that the candidate to be tested must be expressed in this form.

From these conditions the method proposed is one of recursive constructions of majority decision function with respect to the number of variables.

Now our procedure to obtain all the representatives of majority decision functions of n variables can be stated as follows: First choose all of the possible conbinations of two majority decision functions M and N such that for Mx_1+N condition 5.1. can be defined and then determine whether the functions Mx_1+N are really majority decision functions by the method of linear programming.

It is apparent that no true representative is lost by the above procedure.

Next the restriction which is imposed upon the combination of M and N for Mx_1+N which must satisfy the condition 5.1. is examined.

From 1) in condition 5.1. it is necessary for both M and N to be positive in $x_2, x_3, \cdots \cdots, x_n$.

2) in condition 5.1. it is required that for M and N the relation

$$x_2 \gtrsim x_3 \cdots\cdots \gtrsim x_n \qquad (16)$$

must hold.

Moreover, for $r_1 \gtrsim x_2$. it is necessary that Mx_1+N satisfy the conditions:

$$\left.\begin{aligned} & m_2 \supseteq n_1 \\ & m_2 = M(0, x_3, \cdots\cdots, x_n) \\ & n_1 = N(1, x_3, \cdots\cdots, x_n). \end{aligned}\right\} \qquad (17)$$

As 5) in section 3 guarantees that the binary relation is a partial ordering, it can be concluded from the above two relations, (16) and (17), that

$$x_1 \gtrsim x_2 \cdots\cdots \gtrsim x_n \qquad (18)$$

From 3) in condition 5.1. and with 4) in section 3 it is necessary that

$$f = Mx_1 + N \subseteq f^* = N^*x_1 + M^*;$$

that is

$$M \subseteq N^*. \qquad (19)$$

Summing up the above requirements for M and N the following conditions are obtained.

Condition 6.1.

1) Both M and N are positive majority decision function of variables, $x_2, \cdots\cdots, x_n$
2) For both M and N
$$x_2 \gtrsim x^3 \cdots\cdots \gtrsim x_n$$
3) $m_2 \supseteq n_1$
4) $M \subset N^*$.

Choosing those combinations of M and N which satisfy condition 6.1., and forming Mx_1+N, we shall solve the following problem.

Problem 6.1.

Under the restriction of the reduced system of inequalities for Mx_1+N, find a set of $[w_1, \cdots\cdots w_n, T]$ which minimizes T.

If Mx_1+N for which Problem 6.1. has a finite solution, it is a canonical representative of an equivalent class and the solution is an optimum realization of Mx_1+N.

If Problem 6.1. has no solution, then Mx_1+N is not a majority decision function.

Solving Problem 6.1. for all combinations of M and N which satisfy Condition 6.1., all the canonical representatives can be obtained.

M and N must be picked up from all the majority decision functions of $(n-1)$ variables which satisfy 1) and 2) in Condition 6.1.. As is seen in section 5, such functions are canonical representatives of $(n-1)$ variables and their dual functions. Hence once the canonical representatives of $(n-1)$ variables

Table 2

NUMBER OF MAJORITY DECISION FUNCTIONS OF n VARIABLES

n	2^{2^n}	Number of equivalent classes of majority decision functions of n variables	Number of majority decision functions of n variables	Number of eqnivalent classes of self-dual majority decision function of n variables
1	4	1	2	1
2	16	1	8	0
3	256	3	72	1
4	65 536	9	1 536	1
5	4 294 967 296	48	86 080	4
6	18 446 774 073 709 551 616	504	1 448 70 40	14

are obtained, all the candidates of M and N for majority decision functions of the variable are easily determined.

The canonical representatives of majority decision functions of one variable are 0 and x, so starting from these functions we can obtain all the canonical representatives of any number of variables recursively.

7. Majority Decision of Six Variables

Following the procedure described in section 6, a program was written for the parametron computer MUSASINO-1 and all the canonical representatives of six or less variables have been obtained. Problem 6.1. was solved by the dual simplex method in that program.

Majority decision functions of five or less variables are identical to those obtained by S. Muroga. Those functions of six variables without vacuous variables are shown in Table 1. The number of equivalent classes are 504 and each function is numbered from 0 to 503 according to the magnitude of $V=\sum_{i=1}^{6} W_i$, which is expected to denote the complexity of the functions to some extent.

The function is conventionally expressed by denoting the input variables by their subscripts. For instance, 12+13+23 denotes the function $x_1x_2+x_1x_3+x_2x_3$.

In the same entry of the table the optimum structure $(w_1, \cdots\cdots, w_6;\ T)$ of the function is shown. To establish the threshold T by putting an appropriate constant input, the constant of 0 must be coupled with a winding of $2T-V-1$ turns. Dual functions can be realized by merely reversing the polarity of the constant input, that is, by coupling the constant of 1 with the same winding.

8. The Number of Majority Decision Functions of Six Variables

In Table 1 only the canonical representatives are shown.

By computing the number of the members of each equivalent class the total number of majority decision functions of six variables is

Table 3

MAXIMUM VALUES OF PARAMETRES OF OPTIMUM MAJORITY DECISION ELEMENTS

n	w_i	K	T	V
2	1	2	2	3
3	2	4	3	5
4	3	8	5	9
5	5	16	9	17
6	9	33	18	35

obtained and is shown in Table 2 with the numbers of majority decision functions of five or less variables which are quoted from ref. (5).

9. Remarks on the result

Some remarks are added here concerning the complete list of majority decision functions of six variables.

First it is remarkable that condition 6.1. is sufficient for majority decision functions of six or less variables. Hence all of the majority decision functions were obtained without fruitless computation.

Second it is interesting that the optimum structures $[w_1, w_2, \cdots\cdots, w_6, T]$ are all integer-valued in spite of the fact that the values of w's and T are arbitrarily normalized as described in Section 4.

The third remark concerns the structure of the solution space of inequalities (14). It has been noted that for maiority decision functions of five or less vsriables the solution space of each majority decision functions is a pointed cone. That is, denoting the system of inequalities as

$$\boldsymbol{Ax}\geq\boldsymbol{b}, \qquad \boldsymbol{x}=\begin{pmatrix} w_1 \\ \vdots \\ w_n \\ T \end{pmatrix} \tag{20}$$

there exists a certain point $\boldsymbol{x}_0$ such that

$$\boldsymbol{Ax}_0\geq\boldsymbol{b} \tag{21}$$

and any solution $\boldsymbol{x}$ can be written as

$$\boldsymbol{x}=\boldsymbol{x}_0+\boldsymbol{x}', \qquad \boldsymbol{Ax}'\geq 0. \tag{22}$$

This means the solution of (14) is a cone with $\boldsymbol{x}_0$ as the sole vertex. From this it is easily shown that the solution of Problem 4.1. must be $\boldsymbol{x}_0$. Moreover, this suggests that $\boldsymbol{x}_0$ is the solution of Problem 4.1. not only for K but also for any target function, if it is linear in W's and T, and if the problem has a finite solution. This means that the structure obtained is optimum in a wider sense than in the sense of the minimum number of turns.

The structures of majority decision functions of six variables have been examined and it has been noted that almost all the majority decision functions have a solution space of a pointed cone but that 15 out of 504 functions have a space with non cone structure. These 15 functions are marked with * in Table 1.

Fourth, some maximum numbers of the optimum total coupling weights $K=2T-1$ reaches 35. The maximum values of V and T are 33 and 18 respectively.

10. Conclusions

Summing up the results obtained:

1) it is shown that the canonical representative of an equivalent class of majority decision functions can be naturally defined,

2) an efficient procedure to obtain majority decision functions is given,

3) all the majority decision functions of six variables are tabulated with their optimum realizations,

4) the total number of majority decision functions of six variables is calculated,

5) Some properties induced from the compliation of majority decision functions are described.

There still remains the problem to investigate theoretically the properties observed empirically in this study.

References

(1) S. Muroga: Logical elements based on majority decision principle and their circuit complexity, *J. of Institute of Electrical Communication Engineers of Japan*, Vol. **42**, 11, p. 993, 1959.

(2) S. Muroga and I. Toda: Theory of majority decision elements, *J. of Institute of Electrical Communication Engineers of Japan*, Vol. **43**, 10, p. 1071, 1960.

(3) S. Muroga: A computer program to find Boolean functions representable by a single logical element based on a majority decision principle, *Electrical Communication Laboratory Technical Journal*, Vol. **8**, 6, p. 614, 1959.

(4) D. G. Willis: Plastic neurons as memory elements, 1959 *I.R.E. Wescon Conv. Rec.*, pt. 4, p. 55, 1959.

(5) S. Muroga, S. Takasu, and I. Toda: Determination of the structure of a majority decision element by the method of linear programming, *J. of Institute of Electrical Communication Engineers of Japan*, Vol. **43**, 12, p. 1408, 1960.

(6) D. Slepian: On the number of symmetry types of Boolean functions of *n* variables, *Canadian Journ. of Math.*, Vol. **5**, 2, p. 185, 1953.

Table 1

MAJORITY-DECISION-FUNCTION OF SIX VARIABLES WITH OPTIMUM STRUCTURE

Number	V	$w_1 - w_6$	T	$f(x_1, x_2, x_3, x_4, x_5, x_6)$
0	6	111111	4	1234+1235+1245+1345+1236+1346+1246+1256+1356+1456+2345+2346+2356+2456+3456
1	6	111111	5	12345+12346+12356+12456+13456+23456
2	6	111111	6	123456
3	7	211111	4	123+124+134+125+135+145+126+136+146+156+2345+2346+2356+2456+3456
4	7	211111	5	1234+1235+1245+1345+1236+1246+1346+1256+1356+1456+23456
5	7	211111	6	12345+12346+12356+12456+13456
6	8	221111	5	123+124+125+1345+126+1346+1356+1456+2345+2346+2356+2456
7	8	221111	6	1234+1235+1245+1236+1246+1256+13456+23456
8	8	221111	7	12345+12346+12356+12456
9	8	311111	5	123+124+134+125+135+145+126+136+146+156+23456
10	8	311111	6	1234+1235+1245+1345+1236+1246+1346+1256+1356+1456
11	9	222111	5	123+124+134+125+135+126+136+1456+234+235+236+2456+3456
12	9	222111	6	123+1245+1345+1246+1346+1256+1356+2345+2346+2356
13	9	222111	7	1234+1235+1236+12456+13456+23456
14	9	222111	8	12345+12346+12356
15	9	321111	5	12+134+135+145+136+146+156+2345+2346+2356+2456
16	9	321111	6	123+124+125+1345+126+1346+1356+1456+23456
17	9	321111	7	1234+1235+1245+1236+1246+1256+13456
18	9	411111	5	12+13+14+15+16+23456
19	9	411111	6	123+124+134+125+135+145+126+136+146+156
20	10	222211	6	123+124+134+1256+1356+1456+234+2356+2456+3456
21	10	222211	7	1234+1235+1245+1345+1236+1246+1346+2345+2346
22	10	222211	8	1234+12356+12456+13456+23456
23	10	222211	9	12345+12346

Number	V	$w_1 \sim w_6$	T	$f(x_1, x_2, x_3, x_4, x_5, x_6)$
24	10	322111	6	123+124+134+125+135+126+136+1456+2345+2346+2356
25	10	322111	7	123+1245+1345+1246+1346+1256+1356+23456
26	10	322111	8	1234+1235+1236+12456+13456
27	10	331111	6	12+1345+1346+1356+1456+2345+2346+2356+2456
28	10	331111	7	123+124+125+126+13456+23456
29	10	331111	8	1234+1235+1245+1236+1246+1256
30	10	421111	6	12+134+135+145+136+146+156+23456
31	10	421111	7	123+124+125+1345+126+1346+1356+1456
32	10	511111	6	12+13+14+15+16
33	11	322211	6	123+124+134+125+135+145+126+136+146+234+2356+2456+3456
34	11	322211	7	123+124+134+1256+1356+1456+2345+2346
35	11	322211	8	1234+1235+1245+1345+1236+1246+1346+23456
36	11	322211	9	1234+12356+12456+13456
37	11	332111	6	12+134+135+136+1456+234+235+236+2456
38	11	332111	7	123+124+125+1345+126+1346+1356+2345+2346+2356
39	11	332111	8	123+1245+1246+1256+13456+23456
40	11	332111	9	1234+1235+1236+12456
41	11	422111	6	11+13+145+146+156+2345+2346+2359
42	11	422111	7	123+124+134+125+135+126+136+1456+23456
43	11	422111	8	123+1245+1345+1246+1346+1256+1356
44	11	431111	7	12+1345+1346+1356+1456+23456
45	11	431111	8	123+124+125+126+13456
46	11	521111	7	12+134+135+145+136+146+156
47	12	322221	7	123+124+134+125+145+135+2345+2346+2356+2456+3456
48	12	322221	8	1234+1235+1245+1345+1236+1246+1346+1256+1356+1456+2345
49	12	322221	9	1234+1235+1245+1345+23456
50	12	332211	7	123+124+134+125+126+1356+1456+234+2356+2456
51	12	332211	8	123+124+1345+1346+1256+2345+2346
52	12	332211	9	1234+1235+1245+1236+1246+13456+23456
53	12	332211	10	1234+12356+12456

Number	V	$w_1 \sim w_6$	T	$f(x_1, x_2, x_3, x_4, x_5, x_6)$
54	12	333111	7	123+124+134+125+135+126+136+234+235+236
55	12	333111	9	123+12456+13456+23456
56	12	333111	10	1234+1235+1236
57	12	422211	7	123+124+134+125+135+145+126+136+146+2345+2346
58	12	422211	8	123+124+134+1256+1356+1456+23456
59	12	422211	9	1234+1235+1245+1345+1236+1246+1346
60	12	432111	7	12+134+135+136+1456+2345+2346+2356
61	12	432111	8	123+124+125+1345+126+1346+1356+23456
62	12	432111	9	123+1245+1246+1256+13456
63	12	441111	8	12+13456+23456
64	12	441111	9	123+124+125+126
65	12	522111	7	12+13+145+146+156+23456
66	12	522111	8	123+124+134+125+135+126+136+1456
67	12	531111	8	12+1345+1346+1356+1456
68	13	332221	7	123+124+134+125+135+145+126+234+235+245+3456
69	13	332221	8	123+124+125+1345+1346+1356+1456+2345+2346+2356+2456
70	13	332221	9	1234+1235+1245+1345+1236+1246+1256+2345
71	13	332221	10	1234+1235+1245+13456+23456
72	13	333211	8	123+124+134+1256+1256+234+2356
73	13	333211	9	123+1245+1345+1246+1346+2345+2346
74	13	333211	11	1234+12356
75	13	432211	7	12+134+135+145+136+146+234+2356+2456
76	13	432211	8	123+124+134+125+126+1356+1456+2345+2346
77	13	432211	9	123+124+1345+1346+1256+23456
78	13	432211	10	1234+1235+1245+1236+1246+13456
79	13	433111	7	12+13+1456+234+235+236
80	13	433111	8	123+124+134+125+135+126+136+2345+2346+2356
81	13	433111	10	123+12456+13456
82	13	442111	8	12+1345+1346+1356+2345+2346+2356
83	13	442111	10	123+1245+1246+1256

Number	V	$w_1 \sim w_6$	T	$f(x_1, x_1, x_3, x_4, x_5, x_6)$
84	13	522211	7	12+13+14+156+2345+2346
85	13	522211	8	123+124+134+125+135+145+126+136+146+23456
86	13	522211	9	123+124+134+1256+1356+1456
87	13	532111	8	12+134+135+136+1456+23456
88	13	532111	9	123+124+125+1345+126+1346+1356
89	13	541111	9	12+13456
90	13	622111	8	12+13+145+146+156
91	14	332222	8	123+124+125+1345+126+1346+1356+1456+2345+2346+2356+2456+3456
92	14	332222	9	1234+1235+1236+1245+1246+1256+1345+1346+1356+1456+2345+2346 +2356+2456
93	14	333221	8	123+124+134+125+135+1456+234+235+2456+3456
94	14	333221	10	1234+1235+1245+1345+1236+2345
95	14	333221	11	1234+1235+12456+13456+23456
96	14	432221	8	123+124+134+125+135+145+126+2345+2346+2356+2456
97	14	432221	9	123+124+125+1345+1346+1356+1456+2345
98	14	432221	10	1234+1235+1245+1345+1236+1246+1256+23456
99	14	432221	11	1234+1235+1245+13456
100	14	433211	8	123+124+134+125+135+126+136+1456+234+2356
101	14	433211	9	123+124+134+1256+1356+2345+2346
102	14	433211	10	123+1245+1345+1246+1346+23456
103	14	442211	8	12+134+1356+1456+234+2356+2456
104	14	442211	9	123+124+125+1345+126+1346+2345+2346
105	14	442211	10	123+124+1256+13456+23456
106	14	442211	11	1234+1235+1245+1236+1246
107	14	443111	8	12+134++135+136+234+235+236
108	14	443111	11	123+12456
109	14	522221	8	123+124+134+125+135+145+126+136+146+156+2345
110	14	522221	9	123+124+134+125+145+135+23456
111	14	532211	8	12+134+135+145+136+146+2345+2346
112	14	532211	9	123+124+134+125+126+1356+1456+23456
113	14	532211	10	123+124+1345+1346+1256

Number	V	$w_1 \sim w_6$	T	$f(x_1, x_2, x_3, x_4, x_5, x_6)$
114	14	533111	8	12+13+1456+2345+2346+2356
115	14	533111	9	123+124+134+125+135+126+136+23456
116	14	542111	9	12+1345+1346+1356+23456
117	14	622211	8	12+13+14+156+23456
118	14	622211	9	123+124+134+125+135+145+126+136+146
119	14	632111	9	12+134+135+136+1456
120	15	333222	9	123+1245+1345+1246+1346+1256+1356+1456+2345+2346+2356+2456 +3456
121	15	333222	10	1234+1235+1245+1345+1236+1246+1346+1256+1356+2345+2346+2356
122	15	433221	8	123+124+134+125+135+145+126+136+234+235+2456+3456
123	15	433221	9	123+124+134+125+135+1456+2345+2346+2356
124	15	433221	10	123+1245+1345+1246+1346+1256+1356+2345
125	15	433221	11	1234+1235+1245+1345+1236+23456
126	15	433221	12	1234+1235+12456+13456
127	15	433311	9	123+124+134+1256+1356+1456+234
128	15	433311	10	123+124+134+2345+2346
129	15	443211	9	123+124+134+125+126+1356+234+2356
130	15	443211	11	123+1245+1246+13456+23456
131	15	532221	9	123+124+134+125+135+145+126+2345
132	15	532221	10	123+124+125+1345+1346+1356+1456+23456
133	15	532221	11	1234+1235+1245+1345+1236+1246+1256
134	15	533211	8	12+13+145+146+234+2356
135	15	533211	9	123+124+134+125+135+126+136+1456+2345+2346
136	15	533211	10	123+124+134+1256+1356+23456
137	15	533211	11	123+1245+1345+1246+1346
138	15	542211	9	12+134+1356+1456+2345+2346
139	15	542211	10	123+124+125+1345+126+1346+23456
140	15	542211	11	123+124+1256+13456
141	15	543111	9	12+134+135+136+2345+2346+2356
142	15	632211	9	12+134+135+145+136+146+23456
143	15	632211	10	123+124+134+125+126+1356+1456

Number	V	$w_1 \sim w_6$	T	$f(x_1, x_2, x_3, x_4, x_5, x_6)$
144	15	633111	9	12+13+1456+23456
145	15	633111	10	123+124+134+125+135+126+136
146	15	642111	10	12+1345+1346+1356
147	15	722211	9	12+13+14+156
148	16	433222	9	123+124+134+125+135+126+136+1456+2345+2346+2356+2456+3456
149	16	433222	10	123+1245+1345+1246+1346+1256+1356+1456+2345+2346+2356
150*	16	433222	11	1234+1235+1245+1345+1236+1246+1346+1256+1356+23456
151	16	433321	9	123+124+134+125+145+135+234+2356+2456+3456
152	16	433321	11	1234+1235+1245+1345+1236+1246+1346+2345
153	16	443221	9	123+124+134+125+135+126+1456+234+235+2456
154	16	443221	10	123+124+125+1345+1346+1356+2345+2346+2356
155	16	443221	11	123+1245+1345+1246+1256+2345
156	16	443221	12	1234+1235+1245+1236+13456+23456
157	16	443221	13	1234+1235+12456
158	16	443311	10	123+124+134+1256+234
159	16	443311	11	123+124+1345+1346+2345+2346
160	16	532222	9	123+124+134+125+135+145+126+136+146+156+2345+2346+2356+2456
161	16	533221	9	123+124+134+125+135+145+126+136+2345+2346+2356
162	16	533221	10	123+124+134+125+135+1456+2345
163	16	533221	12	1234+1235+1245+1345+1236
164	16	533311	9	123+124+134+125+135+145+126+136+146+234
165	16	533311	11	123+124+134+23456
166	16	542221	9	12+124+135+145+2345+2346+2356+2456
167	16	542221	10	123+134+125+1345+126+1346+1356+1456+2345
168	16	542221	11	123+124+125+1345+23456
169	16	543211	9	12+134+135+136+1456+234+2356
170	16	543211	10	123+124+134+125+126+1356+2345+2346
171	16	543211	12	123+1245+1246+13456
172	16	544111	9	12+13+234+235+236
173	16	552211	10	12+1345+1346+2345+2346

Number	V	$w_1 \sim w_6$	T	$f(x_1, x_2, x_3, x_4, x_5, x_6)$
174	16	552211	12	123+123+1256
175	16	632221	9	12+134+135+145+136+146+156+2345
176	16	632221	10	123+124+134+125+135+145+126+23456
177	16	632221	11	123+124+125+1345+1346+1356+1456
178	16	633211	9	12+13+145+146+2345+2346
179	16	633211	11	123+124+134+1256+1356
180	16	642211	10	12+134+1356+1456+23456
181	16	642211	11	123+124+125+1345+126+1346
182	16	643111	10	12+134+135+136+23456
183	16	722221	9	12+13+14+15+23456
184	16	732211	10	12+134+135+145+136+146
185	16	733111	10	12+13+1456
186	17	433322	10	123+124+134+1256+1356+1456+2345+2346+2356+2456+3456
187	17	433322	11	1234+1235+1245+1345+1236+1246+1346+1256+1356+1456+2345+2346
188	17	443321	10	123+124+125+134+1356+1456+234+2356+2456
189	17	443321	12	1234+1235+1236+1246+1245+1345+2345
190*	17	533222	11	123+1245+1345+1246+1346+1256+1356+1456+23456
191	17	533321	10	123+124+134+125+145+135+2345+2346
192	17	533321	11	123+124+134+1256+1356+1456+2345
193	17	543221	9	12+134+135+145+136+234+235+2456
194	17	543221	10	123+124+134+125+135+126+1456+2345+2346+2356
195	17	543221	11	123+124+125+1345+1346+1356+2345
196	17	543221	12	123+1245+1345+1246+1256+23456
197	17	543221	13	1234+1235+1245+1236+13456
198	17	543311	10	123+124+134+125+126+1356+1456+234
199	17	543311	11	123+124+134+1256+2345+2346
200	17	543311	12	123+124+1345+1346+23456
201	17	544211	10	123+124+134+125+135+126+136+234+2356
202	17	552221	11	123+124+125+1345+126+2345
203	17	552221	12	123+124+125+13456+23456

Number	V	$w_1 \sim w_6$	T	$f(x_1, x_2, x_3, x_4, x_5, x_6)$
204	17	553211	10	12+134+1356+234+2356
205	17	553211	13	123+1245+1246
206	17	633221	10	123+124+134+125+135+145+126+136+2345
207	17	633221	11	123+124+134+125+135+1456+23456
208	17	643211	10	12+134+135+136+1456+2345+2346
209	17	643211	11	123+124+134+125+126+1356+23456
210	17	644111	10	12+13+2345+2346+2356
211	17	652211	11	12+1345+1346+23456
212	17	732221	11	123+124+134+125+135+145+126
213	17	733211	10	12+13+145+146+23456
214	17	742211	11	12+134+1356+1456
215	17	743111	11	12+134+135+136
216	18	443322	10	123+124+134+125+126+1356+1456+234+2356+2456+3456
217	18	443322	11	123+124+1256+1345+1346+1356+1456+2345+2346+2356+2456
218	18	443322	12	1234+1235+1236+1245+1246+1256+1345+1346+2345+2346
219	18	443331	10	123+124+134+125+145+135+234+235+245+3456
220	18	533322	10	123+124+134+125+135+145+126+136+146+2345+2346+2356+2456+3456
221	18	543222	10	123+124+134+125+135+126+136+1456+2345+2346+2356+2456
222*	18	543222	11	123+124+125+1345+126+1346+1356+1456+2345+2346+2356
223	18	543321	10	123+124+134+125+135+145+126+234+2356+2456
224	18	543321	11	123+124+125+134+1356+1456+2345+2346
225	18	543321	12	123+124+1345+1346+1256+2345
226	18	543321	13	1234+1235+1236+1246+1245+1345+23456
227	18	544221	10	123+124+134+125+135+126+136+1456+234+235
228	18	544221	11	123+124+134+125+135+2345+2346+2356
229	18	544221	13	123+1245+1345+23456
230	18	544311	11	123+124+134+1256+1356+234
231	18	553221	10	12+134+135+1456+234+235+2456
232	18	553221	14	1234+1235+1236+1245
233	18	553311	11	123+124+125+126+134+234

Number	V	$w_1 \sim w_6$	T	$f(x_1, x_2, x_3, x_4, x_5, x_6)$
234	18	553311	13	123+124+13456+23456
235	18	633222	10	123+124+134+125+135+145+126+136+146+156+2345+2346+2356
236	18	633222	13	1234+1235+1245+1345+1236+1246+1346+1256+1356
237	18	643221	10	12+134+135+145+136+2345+2346+2356
238	18	643221	11	123+124+134+125+135+126+1456+2345
239	18	643221	12	123+124+125+1345+1346+1356+23456
240	18	643221	13	123+1245+1345+1246+1256
241	18	643311	10	12+134+135+145+136+146+234
242	18	643311	12	123+124+134+1256+23456
243	18	643311	13	123+124+1345+1346
244	18	644211	10	12+13+1456+234+2356
245	18	644211	11	123+124+134+125+135+126+136+2345+2346
246	18	652221	11	12+1345+1346+1356+1456+2345
247	18	652221	12	123+124+125+1345+126+23456
248	18	652221	13	123+124+125+13456
249	18	653211	11	12+134+1356+2345+2346
250	18	733221	10	12+13+145+146+156+2345
251	18	733221	11	123+124+134+125+135+145+126+136+23456
252	18	733221	12	123+124+134+125+135+1456
253	18	733311	10	12+13+14+2345+2346
254	18	742221	11	12+134+135+145+23456
255	18	743211	12	123+124+134+125+126+1356
256	18	744111	11	12+13+23456
257	18	752211	12	12+1345+1346
258	18	833211	11	12+13+145+146
259	19	443332	11	123+124+125+1345+1346+1356+1456+2345+2346+2356+2456+3456
260*	19	543322	11	123+124+134+125+126+1356+1456+2345+2346+2356+2456
261	19	543322	12	123+124+1256+1345+1346+1356+1456+2345+2346
262	19	543322	13	1234+1235+1236+1245+1246+1256+1345+1346+23456
263	19	543331	11	123+124+134+125+145+135+2345+2346+2356+2456

Number	V	$w_1 \sim w_6$	T	$f(x_1, x_2, x_3, x_4, x_5, x_6)$
264	19	544321	11	123+124+134+125+135+1456+234+2356
265	19	544321	13	123+1245+1345+1246+1346+2345
266	19	553321	12	123+124+125+1345+1346+2345+2346
267	19	553321	13	123+124+1256+1345+2345
268	19	554221	11	123+124+125+126+134+135+234+235
269	19	554221	14	123+1245+13456+23456
270	19	643321	11	123+124+134+125+135+145+126+2345+2346
271	19	643321	12	123+124+125+134+1356+1456+2345
272	19	643321	14	1234+1235+1236+1246+1245+1345
273	19	644311	11	123+124+134+125+135+126+136+1456+234
274	19	653221	11	12+134+135+1456+2345+2346+2356
275	19	653221	12	123+124+135+1345+126+1346+1356+2345
276	19	653311	11	12+134+1356+1456+234
277	19	653311	12	123+124+125+126+134+2345+2346
278	19	653311	14	123+124+13456
279	19	654211	11	12+134+135+136+234+2356
280	19	733222	13	123+1245+1345+1246+1346+1256+1356+1456
281	19	733321	11	123+124+134+125+135+145+126+136+146+2345
282	19	743221	11	12+134+135+145+136+2345
283	19	743221	12	123+124+134+125+135+126+1456+23456
284	19	743221	13	123+124+125+1345+1346+1356
285	19	743311	13	123+124+134+1256
286	19	744211	11	12+13+1456+2345+2346
287	19	752221	13	123+124+125+1345+126
288	19	753211	12	12+134+1356+23456
289	19	833221	12	123+124+134+125+145+126+136
290	19	833311	11	12+13+14+23456
291	20	543332	11	123+124+134+125+135+145+126+2345+2346+2356+2456+3456
292	20	544322	11	123+124+134+125+135+126+136+1456+234+2356+2456+3456
293	20	544322	12	123+124+134+1256+1356+1456+2345+2346+2356

Number	V	$w_1 \sim w_6$	T	$f(x_1, x_2, x_3, x_4, x_5, x_6)$
294	20	544322	13	123+1245+1345+1246+1346+1256+1356+2345+2346
295	20	544331	11	123+124+134+125+145+135+234+235+2456+3456
296	20	553331	11	123+124+134+125+135+145+126+234+235+245
297	20	554222	11	123+124+134+125+135+126+136+1456+234+235+236+2456
298	20	554321	12	123+124+125+134+1356+234+2356
299	20	554321	14	123+1245+1246+1345+2345
300	20	643322	11	123+124+134+125+135+145+126+136+146+2345+2346+2356+2456
301*	20	643322	13	123+124+1256+1345+1346+1356+1456+23456
302	20	644321	11	123+124+134+125+135+145+126+136+234+2356
303	20	644321	12	123+124+134+125+135+1456+2345+2346
304	20	644321	13	123+124+134+1256+1356+2345
305	20	653222	11	12+134+135+136+1456+2345+2346+2356+2456
306	20	653321	11	12+134+135+145+234+2356+2456
307	20	653321	13	123+124+125+1345+1346+2345
308	20	653321	14	124+124+1256+1345+23456
309	20	654221	11	12+134+135+136+1456+234+235
310	20	654221	12	123+124+125+126+134+135+2345+2346+2356
311	20	654221	15	123+1245+13456
312	20	654311	12	123+124+134+125+126+1356+234
313	20	733322	11	123+124+134+125+135+145+126+136+146+156+2345+2346
314	20	743222	11	12+134+135+145+136+146+156+2345+2346+2356
315	20	743321	13	123+124+125+134+1356+1456+23456
316	20	744221	11	12+13+145+2345+2346+2356
317	20	744221	12	123+124+134+125+135+126+136+1456+2345
318	20	744221	13	123+124+134+125+135+23456
319	20	744311	11	12+13+145+146+234
320	20	753221	12	12+134+135+1456+2345
321	20	753311	13	123+124+125+126+134+23456
322	20	754211	12	12+134+135+136+2345+2346
323	20	762221	13	12+1345+23456

Number	V	$w_1 \sim w_6$	T	$f(x_1, x_2, x_3, x_4. x_5, x_6)$
324	20	833321	11	12+13+14+156+2345
325	20	843221	12	12+134+135+145+136+23456
326	20	843221	13	123+124+134+125+135+126+1456
327	20	853211	13	12+134+1356
328	21	544332	12	123+124+134+125+135+1456+2345+2346+2356+2456+3456
329	21	544332	14	1234+1235+1245+1345+1236+1246+1346+1256+1356+2345
330	21	554322	13	123+124+1256+1345+1356+1346+2345+2346+2356
331	21	554322	14	123+1245+1256+1246+1345+1346+2345+2346
332	21	554331	12	123+124+134+125+135+1456+234+235+2456
333	21	644331	12	123+124+134+125+145+135+2345+2346+2356
334*	21	653322	13	123+124+125+1345+126+1346+1356+1456+2345+2346
335	21	654321	12	123+124+134+125+135+126+1456+234+2356
336	21	654321	13	123+124+125+134+1356+2345+2346
337	21	654321	15	123+1245+1246+1345+23456
338	21	655221	12	123+124+134+125+135+126+136+234+235
339	21	743322	15	1234+1235+1236+1245+1246+1256+1345+1346
340	21	744321	12	123+124+134+125+135+145+126+136+2345+2346
341	21	753321	12	12+134+135+145+2345+2346
342	21	753321	13	123+124+134+125+126+1356+1456+2345
343	21	753321	14	123+124+125+1345+1346+23456
344	21	753321	15	123+124+1256+1345
345	21	754221	13	123+124+125+126+134+135+2345
346	21	754311	12	12+134+135+ +136+1456+234
347	21	755211	12	12+13+234+2356
348	21	763221	13	12+1345+1346+1356+2345
349	21	763311	13	12+134+2345+2346
350	21	843321	12	12+134+135+145+136+146+2345
351	21	843321	14	123+124+125+134+1356+1456
352	21	853221	13	12+134+135+1456+23456
353	21	853311	14	123+124+125+126+134

Number	V	$w_1 \sim w_6$	T	$f(x_1, x_2, x_3, x_4, x_5, x_6)$
354	21	943221	13	12+134+135+145+136
355	22	544333	14	1234+1235+1245+1345+1236+1246+1346+1256+1356+1456+2345+2346+2356
356	22	554332	12	123+124+134+125+135+126+1456+234+235+2456+3456
357	22	554422	13	123+124+134+1256+1356+1456+234+2356+2456
358	22	644332	12	123+124+134+125+135+145+126+136+2345+2346+2356+2456+3456
359	22	644332	14	123+1245+1345+1246+1346+1256+1356+1456+2345
360	22	654322	12	123+124+134+125+135+126+136+1456+234+2356+2456
361*	22	654322	13	123+124+134+125+126+1356+1456+2345+2346+2356
362	22	654322	14	123+124+1256+1345+1356+1346+2345+2346
363*	22	654322	15	123+1245+1256+1246+1345+1346+23456
364	22	654331	12	123+124+134+125+135+145+126+234+235+2456
365	22	654421	13	123+124+125+134+1356+1456+234
366	22	654421	15	123+124+1345+1346+2345
367	22	655222	12	123+124+134+125+135+126+136+1456+234+235+236
368	22	655321	13	123+124+134+125+135+234+2356
369	22	744322	12	123+124+134+125+135+145+126+136+146+2345+2346+2356
370	22	753322	12	12+134+135+145+136+146+2345+2346+2356+2456
371	22	754321	12	12+134+135+145+136+234+2356
372	22	754321	13	123+124+134+125+135+126+1456+2345+2346
373	22	754321	14	123+124+125+134+1356+2345
374	22	754321	16	123+1245+1246+1345
375	22	755221	12	12+13+1456+234+235
376	22	755311	13	123+124+134+125+135+126+136+234
377	22	763222	13	12+1345+1346+1356+1456+2345+2346+2356
378	22	763321	14	123+124+125+1345+126+1346+2345
379	22	764221	13	12+134+135+2345+2346+2356
380	22	764311	13	12+134+1356+234
381	22	843322	12	12+134+135+145+136+146+156+2345+2346
382	22	843322	15	123+124+1256+1345+1346+1356+1456
383	22	853321	15	123+124+125+1345+1346

Number	V	$w_1 \sim w_6$	T	$f(x_1, x_2, x_3, x_4, x_5, x_6)$
384	22	854221	13	12+134+135+136+1456+2345
385	22	854221	14	123+124+125+126+134+135+23456
386	22	855211	13	12+13+2345+2346
387	22	863311	14	12+134+23456
388	22	944221	13	12+13+145+23456
389	22	953221	14	12+134+135+1456
390	23	554333	14	123+1245+1345+1246+1346+1256+1356+1456+2345+2346+2356+2456
391	23	554432	13	123+124+125+134+1356+1456+234+2356+2456+3456
392*	23	654332	13	123+124+134+125+135+126+1456+2345+2346+2356+2456
393	23	654332	14	123+124+125+1345+1346+1356+1456+2345+2346+2356
394	23	654431	13	123+124+134+125+145+135+234+2356+2456
395	23	655322	14	123+124+134+1256+1356+2345+2346+2356
396	23	655331	13	123+124+134+125+135+1456+234+235
397*	23	754322	15	123+124+1256+1345+1356+1346+23456
398	23	754331	13	123+124+134+125+135+145+126+2345+2346+2356
399	23	754421	13	123+124+134+125+135+145+126+234
400	23	754421	15	123+124+134+1256+2345
401	23	755321	14	123+124+134+125+135+2345+2346
402	23	763322	13	12+134+1356+1456+2345+2346+2356+2456
403	23	764321	13	12+134+135+1456+234+2356
404	23	765221	13	12+134+135+136+234+235
405	23	854321	13	12+134+135+145+136+2345+2346
406	23	854321	15	123+124+125+134+1356+23456
407	23	855221	14	123+124+134+125+135+126+136+2345
408	23	855311	13	12+13+1456+234
409	23	863321	14	12+134+1356+1456+2345
410	23	944321	13	12+13+145+146+2345
411	23	954221	15	123+124+125+126+134+135
412	24	554433	14	123+124+1256+1345+1346+1356+1456+2345+2346+2356+2456+3456
413	24	654432	13	123+124+134+125+135+145+126+234+2356+2456+3456

Number	V	$w_1 \sim w_2$	T	$f(x_1, x_2, x_3, x_4, x_5, x_6)$
414	24	654432	14	123+124+125+134+1356+1456+2345+2346+2356+2456
415	24	654432	16	1234+1235+1236+1245+1246+1256+1345+1346+2345
416	24	655332	13	123+124+134+125+135+126+136+1456+234+235+2456+3456
417	24	655422	14	123+124+134+1256+1356+1456+234+2356
418	24	754332	13	123+124+134+125+135+145+126+136+2345+2346+2356+2456
419	24	754422	13	123+124+134+125+135+145+126+136+146+234+2356+2456
420	24	755331	13	123+124+134+125+135+145+126+136+234+235
421	24	755421	14	123+124+134+125+135+1456+234
422	24	764322	13	12+134+135+136+1456+234+2356+2456
423*	24	764322	15	123+124+125+1345+126+1346+1356+2345+2346
424	24	764331	13	12+134+135+145+234+235+2456
425	24	764421	15	123+124+134+125+2345+2346
426	24	764421	17	123+124+1345+23456
427	24	765222	13	12+134+135+136+1456+234+235+236
428	24	765321	14	123+124+125+126+134+135+234+2356
429	24	854322	13	12+134+135+145+136+146+2345+2346+2356
430	24	854322	17	123+1245+1256+1246+1345+1346
431	24	855321	13	12+13+145+234+2356
432	24	864321	14	12+134+135+1456+2345+2346
433	24	864321	15	123+124+134+125+126+1356+2345
434	24	865311	14	12+134+135+136+234
435	24	873321	15	12+1345+1346+2345
436	24	944322	13	12+13+145+146+156+2345+2346
437	24	954321	16	123+124+125+134+1356
438	24	955221	14	12+13+1456+2345
439	24	964221	15	12+134+135+23456
440	25	654433	14	123+124+134+125+126+1356+1456+2345+2346+2356+2456+3456
441	25	655432	14	123+124+134+125+135+1456+234+2356+2456+3456
442	25	754432	16	123+124+1256+1345+1346+1356+1456+2345
443	25	755431	14	123+124+134+125+145+135+234+2356

Number	V	$w_1 \sim w_6$	T	$f(x_1, x_2, x_3, x_4, x_5, x_6)$
444*	25	765322	15	123+124+134+125+126+1356+2345+2346+2356
445	25	765331	14	123+124+134+125+135+126+1456+234+235
446	25	765421	15	123+124+125+1356+234
447	25	855421	14	123+124+134+125+135+145+126+136+234
448	25	864331	14	12+134+135+145+2345+2346+2356
449	25	865321	15	123+124+125+126+134+135+2345+2346
450	25	873322	15	12+1345+1346+1356+1456+2345+2346
451	25	954322	17	123+124+1256+1345+1356+1346
452	25	955321	14	12+13+145+2345+2346
453	25	965221	15	12+134+135+136+2345
454	26	655433	17	1234+1235+1245+1345+1236+1246+1346+1256+1356+2345+2346
455	26	755432	14	123+124+134+125+135+145+126+136+234+2356+2456+3456
456	26	764432	16	123+124+125+1345+1346+1356+1456+2345+2346
457	26	765332	14	123+124+134+125+135+126+136+1456+234+235+2456
458*	26	765422	15	123+124+134+125+126+1356+1456+234+2356
459	26	855422	14	123+124+134+125+135+145+126+136+146+234+2356
460	26	864332	14	12+134+135+145+136+2345+2346+2356+2456
461	26	865331	14	12+134+135+145+136+234+235
462	26	865421	15	123+124+134+125+135+126+1456+234
463	26	874322	15	12+134+1356+1456+2345+2346+2356
464	26	875321	15	12+134+135+234+2356
465	26	955322	14	12+13+145+146+2345+2346+2356
466	26	964421	17	123+124+134+125+23456
467	26	974321	16	12+134+1356+2345
468	27	755433	17	123+1245+1345+1246+1346+1256+1356+1456+2345+2346
469*	27	765432	15	123+124+134+125+135+126+1456+234+2356+2456
470	27	765432	16	123+124+125+134+1356+1456+2345+2346+2356
471	27	765432	18	123+1245+1256+1246+1345+1346+2345
472	27	765441	15	123+124+134+125+145+135+234+235+2456
473	27	865431	15	123+124+134+125+135+145+126+234+2356

Number	V	$w_1 \sim w_6$	T	$f(x_1, x_2, x_3, x_4, x_5, x_6)$
474	27	874332	15	12+134+135+1456+2345+2346+2356+2456
475	27	875331	15	12+134+135+1456+234+235
476	27	965421	15	12+134+135+145+136+234
477	27	974421	17	123+124+125+126+134+2345
478	27	975321	16	12+134+135+2345+2346
479	28	765433	17	123+124+1256+1345+1346+1356+1456+2345+2346+2356
480	28	765442	15	123+124+134+125+135+145+126+234+235+2456+3456
481	28	765442	16	123+124+134+125+135+1456+2345+2346+2356+2456
482	28	865432	15	123+124+134+125+135+145+126+136+234+2356+2456
483	28	865432	18	123+124+1256+1345+1356+1346+2345
484	28	875332	15	12+134+135+136+1456+234+235+2456
485	28	965422	15	12+134+135+145+136+146+234+2356
486	28	975421	16	12+134+135+1456+234
487	28	984322	17	12+1345+1346+1356+2345+2346
488	29	765443	16	123+124+134+125+135+126+1456+2345+2346+2356+2456+3456
489	29	765533	17	123+124+134+1256+1356+1456+2345+2346+2356+2456
490	29	875432	18	123+124+125+1345+1346+1356+2345+2346
491	29	975431	16	12+134+135+145+234+2356
492	29	976421	17	123+124+125+126+134+135+234
493	29	985322	17	12+134+1356+2345+2346+2356
494	30	765543	17	123+124+125+134+1356+1456+2345+2346+2356+2456+3456
495	30	876432	18	123+124+125+134+1356+2345+2346+2356
496	30	975432	16	12+134+135+145+136+234+2356+2456
497	30	985422	17	12+134+1356+1456+234+2356
498	31	876532	18	123+124+125+134+1356+1456+234+2356
499	31	976441	17	123+124+134+125+135+145+126+234+235
500	31	985432	17	12+134+135+1456+234+2356+2456
501	32	876542	18	123+124+134+125+135+1456+234+2356+2456
502	32	976442	17	123+124+134+125+135+145+126+136+234+235+2456
503	33	876543	18	123+124+134+125+135+126+1456+234+2356+2456+3456

REVIEW OF THE ELECTRICAL COMMUNICATION LABORATORY
VOLUME 9, NUMBERS 5-6, MAY-JUNE, 1961

U.D.C. 547. 298 : 546. 824. 36'431] : 537. 226

Ferroelectric Polarization Reversal by Linear Fields*

***Hiroyuki SHIBATA,*†** ***Hiroo TOYODA,*†** and ***Shigeru WAKU*†**

Ferroelectric polarization reversal of triglycine sulfate, and pure and impurity-doped barium titanate crystals was studied by the application of linearly rising field pulses The coercive fields of the substances increase with the dE/dt of the pulses, which ranged from 10^6 to 10^{10} $V\,cm^{-1}sec^{-1}$, and from the results information on the high frequency characteristics up to several hundred kilocycles per second is derived. The value of $d(\log E_c)$ $d(\log dE/dt)$ rises with dE/dt and approaches 1/2. A simplified phenomenological treatment of polarization reversal is presented, and the calculated values are in a fairly good agreement with the experimental results for triglycine sulfate. The effect of the setting condition is briefly discussed.

1. Introduction

Polarization reversal of ferroelectrics depends on the measuring conditions, for example, the waveform of the reversing field and the setting condition; the former is characterized by its dE/dt and the latter means setting voltage, duration, and off-time. For the observation of the hysteresis loop, a continuous sine wave is usually used. The measurement includes, however, some inconveniences. Both dE/dt and the setting conditions of the field depend on amplitude and frequency, and the effects can not be separated from each other. At high frequencies, the observation is disturbed by heating due to hysteresis loss. However, the application of linearly-rising field pulses excludes these faults. The pulses are characterized by a single time-indepent parameter dE/dt and the setting conditions can be varied independently. Frequency dependence of the reversal is readily obtained without disturbance due to a heating loss.

Polarization reversal by a linear field has been performed on GASH crystals by Prutton.(1) He had studied the change of maximum switching current with dE/dt ranging from 10^6 to 10^8 V cm^{-1}sec^{-1}. The authors expanded the range from 10^6 to 10^{10}V cm^{-1} $\cdot$sec^{-1} and studied the non-linear characteristics of triglycine sulfate (hereafter abbreviated TGS) and barium titanate single crystals.

2. Experimental Methods

The pulses have the waveform shown in Fig. 1. The pulse width varies from 1 μ sec to 5 msec and the repetition frequency from 0.1 to 100 c/s. The field strength in the crystals is so chosen that the pulse maximum and the negative setting field are 10 to 15 kV/cm, respectively. The polarization curves are observed on a cathode-ray oscilloscope by using the so-called Sawyer-Tower circuit.(2)

Specimens examined were TGS crystals grown from the aqueous solution, barium titanate crystals grown by the KF flux method; and barium titanate crystals containing Nb, Ta,

---

* MS in Japanese received by the Electrical Communcation Laboratory, Jan. 23, 1961. Originally published in the *Kenkyû Zituyôka Hôkoku* (*Electrical Communication Laboratory Technical Journal*), N.T.T , Vol. 10, No. 6, pp. 1197–1205, 1961.

† Ferroelectric Materials Research Section.

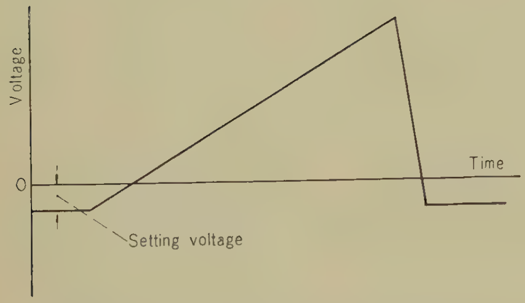


Fig. 1—Waveform of a linear field.

Y, and, V of 0.1 mol per cent. The crystal thickness ranged from 0.002 to 0.02 cm, and gold electrodes with diameter of 0.1 cm were evaporated. Barium titanate crystals were etched in phosphoric acid, and c-domained crystals after poling were used.

3. Experimental Results

The polarization curves against field observed on a cathode-ray oscilloscope correspond to one half of the hysteresis loop. Fig. 2 shows the curves for TGS, taking dE/dt as the parameter. Coercive field E_c is defined as the field where the average polarization vanishes, as in the case of sine wave measurement. Figs. 3 and 4 give the relations between E_c and dE/dt for different temperatures for TGS and BaTiO$_3$, respectively. The gradient of the curves, d(logE_c)/d(logd$E$/dt), increases with dE/dt; and in the case of TGS, it approaches 1/2. The changes agree with the results of phenomenological theory described later. The results obtained for impurity-doped barium titanate are shown in Fig. 5, and are not so different from those obtained for pure crystals. The crystal thickness has such a large effect that the effects of impurities may be masked off.

Repetition frequency has larger effects for TGS than for barium titanate. At room temperature, a frequency increase from 1 to 100 c/s causes a ten per cent decrease of E_c for TGS. A similar decrease is found in barium titanate containing V$_2$O$_5$ and Nb$_2$O$_5$.

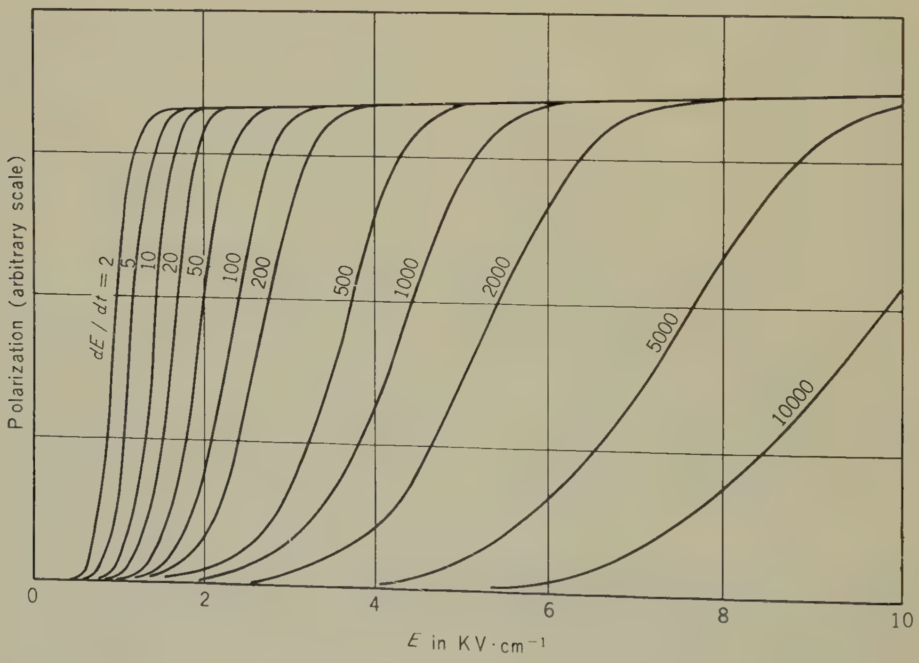


Fig. 2—Polarization curves for TGS. dE/dt is represented in 10^6 V·cm^{-}·sec^{-1}.

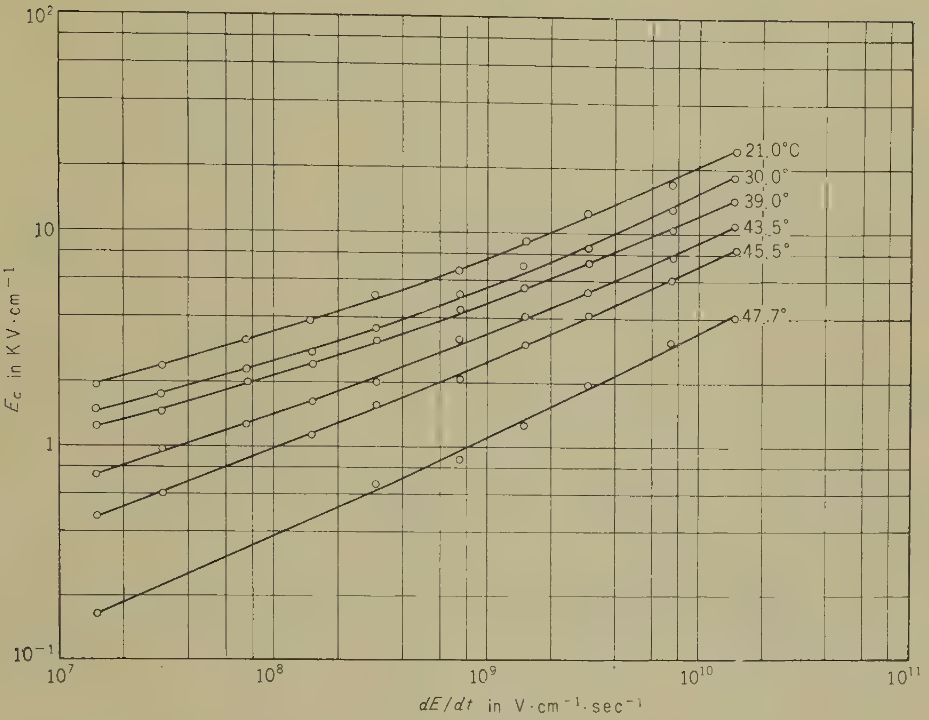


Fig. 3—Coercive field vs. dE/dt curves for TGS at various temperature.

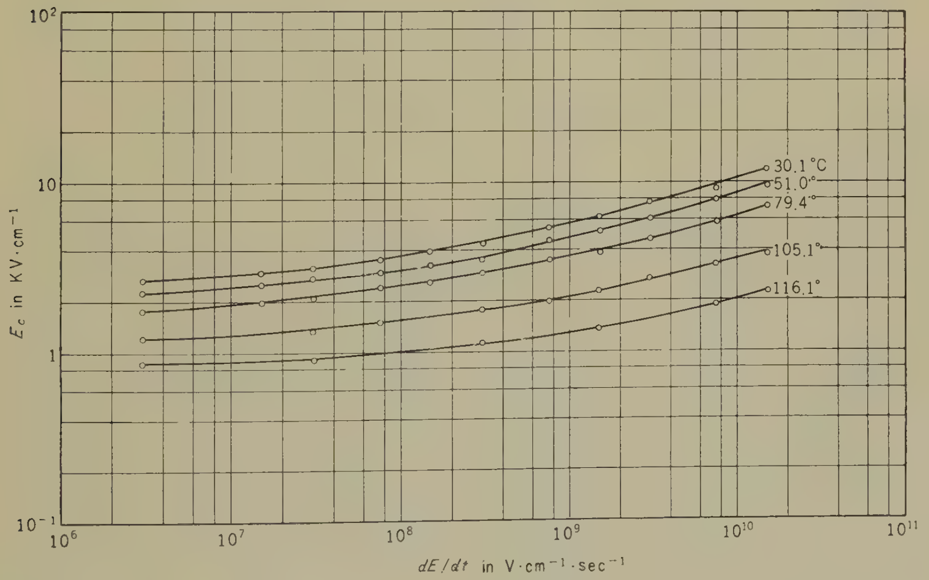


Fig. 4—Coercive field vs. dE/dt curves for $BaTiO_3$ at various temperatures.

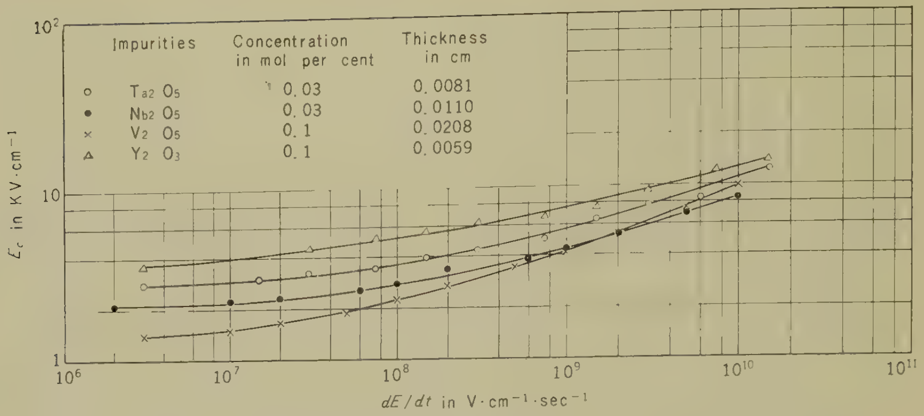
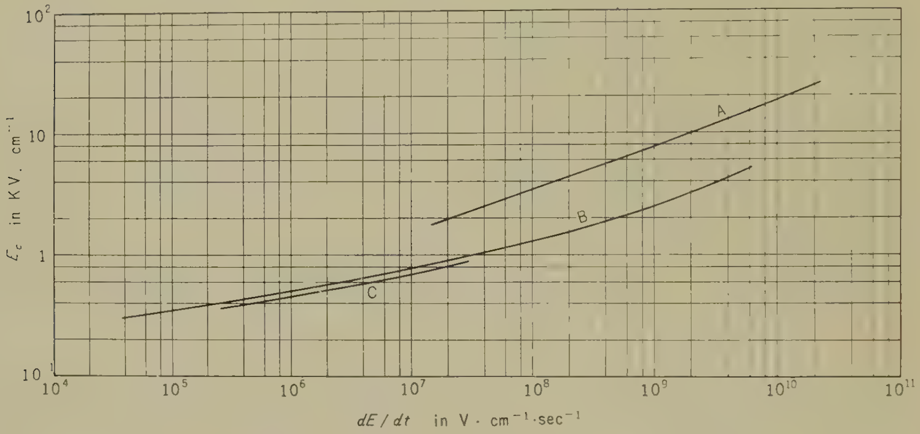
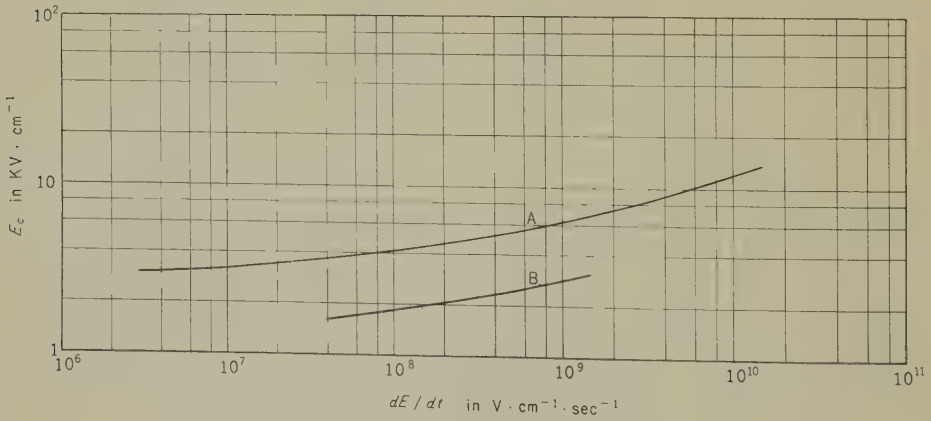


Fig. 5—Coercive field vs. dE/dt for $BaTiO_3$ containing impurities at 30° C.



A: Linear field, B: Sine wave field; according to Pulvari,
C: Sine wave field; according to Domansky.

Fig. 6—Comparison of E_c for TGS obtained by two waveforms at room temperature.



A: Linear field, B: Sine wave field; according to Husimi.

Fig, 7—Comparison of E_c for $BaTiO_3$ obtained by two waveforms at room temperature.

Figs. 2 to 5 show the results of measurements at 2 c/s.

In order to compare the results with those obtained by a sine wave field, the authors will approximate the sine wave as a linear field with dE/dt given as follows

$$\frac{dE}{dt}=2\pi fE_m, \qquad (1)$$

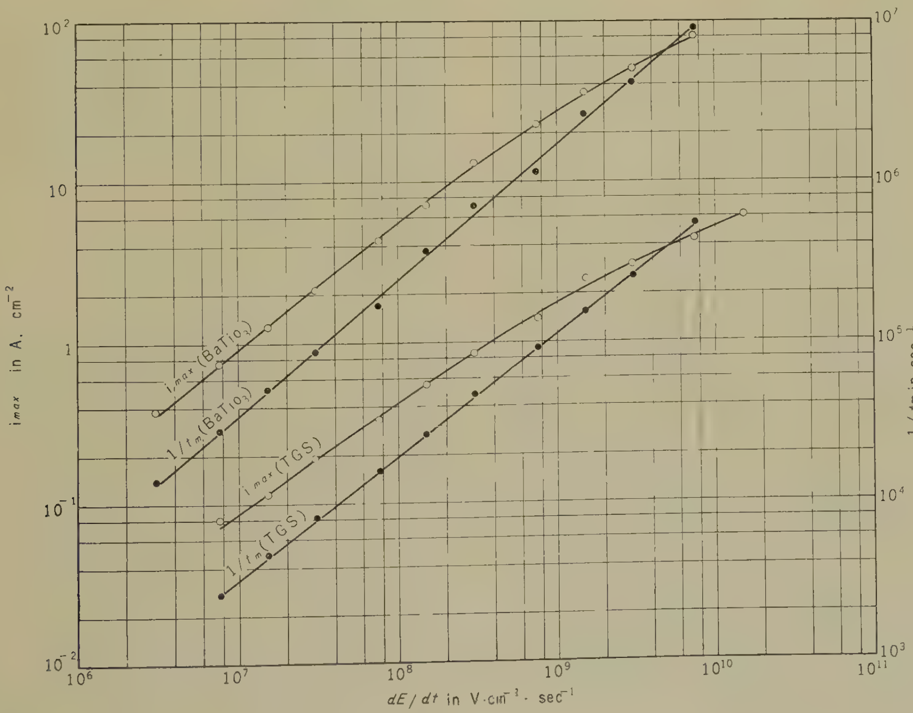
where f and E_m stand for the frequency and amplitude respectively of the sine wave. Relations between E_c and dE/dt converted by the use of Eq. (1) are compared with those obtained by the linear field in Figs. 6 and 7. The sine wave values are cited from the work of Domansky⁽³⁾ and Pulavari⁽⁴⁾ for TGS, and from the work of Husimi* for $BaTiO_3$. Differences are more remarkable for $BaTiO_3$. It is thought that the disagreement rises in setting conditions rather than in the failure of the approximation of Eq. (1), since a continuous wave is used in the sine wave measurement. The problem will be discussed later.

The maximum switching current, i_{max}, and the time, t_m, when the current becomes maximum are shown in Fig. 8. Figs. 6 and 10 give the dependences of the coercive field on the temperature.

4. Discussion

The authors treat the ferroelectric polarization reversal from a phenomenological viewpoint. The following assumptions are first made:

1. Polarization reversal occurs in two steps, that is, domain nucleation and domain growth.
2. All the domain nuclei have a common and constant cross section and they sprout on one side of the crystal.



i_{max}: The maximum switching current, t_m: The time at which the current becomes the maximum.

Fig. 8—Switching current and switching time variations with dE/dt.

---

* Private communication.

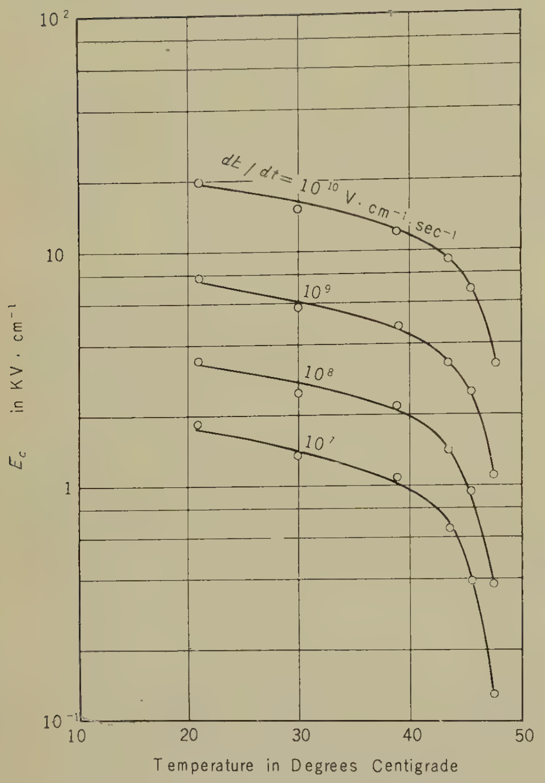


Fig. 9—Temperature dependences of coercive field in TGS.

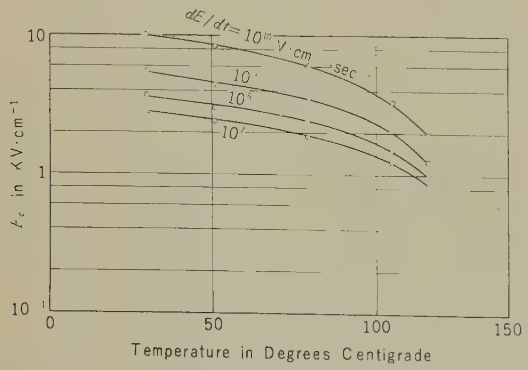


Fig. 10—Temperature dependences of coercive field in BaTiO$_3$.

3. The probability of domain nucleation, p, depends on the applied field.

4. The velocity of domain growth, v, is proportional to the field, that is, $v=\mu(E-E_0)$, μ representing mobility.

The number of remaining nuclei is given by

$$N=N_0 \exp\left\{-\frac{1}{c}\int_0^E p(E)\mathrm{d}E\right\}, \qquad (2)$$

where $c=\mathrm{d}E/\mathrm{d}t$ is constant for a linear field. Then, the following equation gives the switching current for the field E:

$$i=\frac{Q\mu}{cl}(E-E_0)\int_{\{(E-E_0)^2-(2cl/\mu)\}^{1/2}+E_0}^{E} \times \exp\left(-\frac{1}{c}\int_0^E p\cdot\mathrm{d}E\right)p\cdot\mathrm{d}E, \qquad (3)$$

where Q is the total charge and l is the crystal thickness. The lower limit of the integral originates from the condition that the reversed domains, which have penetrated through the crystal, no longer contribute to the current. In the experiment described previously, it is thought that the condition $E-E_0 \gg (2cl/\mu)^{1/2}$ is satisfied. Then, Eq. (3) is reduced as follows;

$$i=Q\exp\left(-\frac{1}{c}\int_0^E p\cdot\mathrm{d}E\right)p. \qquad (4)$$

The current merely depends on the rate of nucleation.

The coercive field E_c and the field E_{max} where the switching current becomes maximum are given by

$$\int_0^{E_c} p\cdot\mathrm{d}E \fallingdotseq 0.7c. \qquad (5)$$

and

$$\frac{1}{c}p^2(E_{max})=\frac{\mathrm{d}p}{\mathrm{d}E}E_{max}, \qquad (6)$$

respectively.

For a most simple dependence of p on E, that is, $p=a'E$, the relation

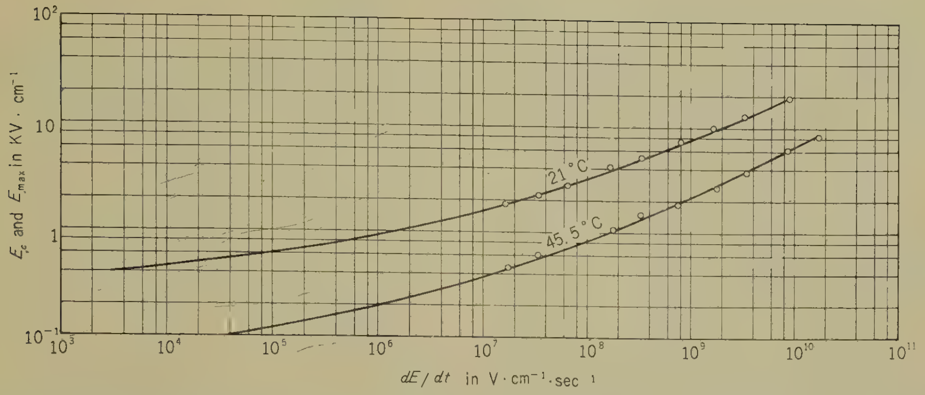


Fig. 11—Comparison of coercive field in TGS. Small circles represent E_c by linear field and solid lines show the curves calculated according to Eq. (8).

$$E_{max}=\frac{E_c}{1.2}=\left(\frac{c}{a'}\right)^{\frac{1}{2}} \tag{7}$$

is obtained. Then $\mathrm{d}(\log E_c)/\mathrm{d}(\log \mathrm{d}E/\mathrm{d}t)$ is equal to 1/2.(5) In a polarization reversal experiment by the use of rectangular pulses, Merz(6) found that the maximum switching current is proportional to $\exp(-\alpha/E)$ for small fields and to E for large fields. If the dependence is dominated by the nucleation rate, it is expected that the relation $p=aE\times\exp(-\alpha/E)$ holds. Substitution into Eq. (6) gives

$$2\log E_{max}-\frac{\alpha}{E_{max}}=\log\frac{c}{a}\left(1+\frac{\alpha}{E_{max}}\right). \tag{8}$$

Eq. (8) expresses that $\mathrm{d}(\log E_{max})/\mathrm{d}(\log \mathrm{d}E/\mathrm{d}t)$ rises with $\mathrm{d}E/\mathrm{d}t$ and approaches 1/2. The deviation of the gradient from 1/2 increases with α. Substituting α(5), obtained from the rectangular experiment in a similar setting condition, into Eq. (8) and neglecting the difference between E_{max} and E_c, the calculated curves are compared with the experimental results in Fig. 11. The agreement is fairly good.

The smaller value of E_c in the continuous sine wave experiment than that in the pulse experiment is thought to correspond to the decrease of α. Fatuzzo(7) and one of the authors(8) have reported the variation of α and the maximum switching current depending on the setting condition. The value of α increases with the negative setting field applied to the crystal and its duration and off-time. The first two effects are probably due to space charge formation, but this is not the case for the third effect.

Irradiated and impurity-doped crystals of TGS exhibit an abnormal hysteresis loop,(9) showing an occurrence of internal biasing field. When the irradiation period is short or the impurity concentration is small, heating-up of or alternating voltage applied to the crystal temporarily removes the anomaly; which reappears, however, with a lapse of time. The results suggest that some rearrangements of bonding around the imperfections arise slowly and have a close relation to the polarization direction. The crystal surface is thought to play the role of such an imperfection and to give rise to a time effect in polarization reversal.

In barium titanate, the effect of the setting condition is less considerable than in TGS and the dependence of E_c on $\mathrm{d}E/\mathrm{d}t$ changes markedly with the crystal thickness. Eq. (3) means that the increase in thickness would invite a higher coercive field. The observed results, however, do not show this to be the case and there should be other effects due to, for example, the surface layer(10) which is now under investigation.

5. Conclusion

The use of a linear field for the polarization reversal experiment of ferroelectrics is very effective to exclude the ambiguities contained in the sine wave measurement and to simplify the analysis of the results. The switching current and accordingly, the coercive field are given by the functions of only one parameter, that is, the time-independent dE/dt of the pulse field, if the setting conditions are kept constant. The coercive field of two ferroelectrics examined increases with dE/dt. In TGS, d(log E_c)/d(log dE/dt) rises with dE/dt and approaches 1/2, independent of temperature. In $BaTiO_3$, the variation with dE/dt is somewhat mild for the same region. A phenomenological consideration of the polarization reversal explains the dependence of the coercive field quite well, according to an assumption that the switching is dominated by the rate of domain nucleation. Coercive field obtained by a linear pulse is larger than by a continuous sine wave with similar dE/dt. The difference is probably due to the dissimilarities of the setting conditions in two waveforms.

The authors wish to express their sincere thanks to Dr. H. Hirabayashi and to Mr. K. Hushimi for their helpful suggestions.

References

(1) M. Prutton; *Proc. Phys, Soc.* **70B**, (1957) 1064. "The Polarization Reversal Process in Guanidine Aluminium Sulphate Hexahydrate."
(2) C. B. Sawyer and C. H. Tower; *Phys, Rev.* **35**, (1930) 269. "Rochelle Salt as a Dielectric."
(3) S. Domansky; *Proc. Phys. Soc.* **72B**, (1958) 306. "The Dependence of the Coercive Field of Triglycine Sulfate on Temperature and Frequency."
(4) C. F. Pulavri and W. Kuebler; *J. Appl. Phys.* **29**, (1958) 1742. "Polarization Reversal in TriGlycine Fluoberyllate and TriGlycine Sulfate Single Crystals."
(5) H. Toyoda, S. Waku, H. Shibata, and Y. Tanaka; *J. Phys. Soc. Japan* **14**, ((1959) 109. "Some Ferroelectrical Properties of Triglycine Sulfate."
(6) W. J. Merz; *Phys. Rev.* **95**, (1954) 690. "Domain Formation and Domain Wall Motion in Ferroelectric $BaTiO_3$ Single Crystals."
(7) E. Fatuzzo; *Helv. Phys. Acta* **31**, (1958) 309. Some Physical Properties of Tri-Glycine Sulfate and its Isomorphs."
(8) H. Toyoda; *Reports of ECL, NTT* **7**, (1959) 416. "Properties of Ferroelectric Glycine Sulfate."
(9) H. Toyoda and Y. Tanaka; *J. Appl. Phys. Japan* **29**, (1960) 521. "Growth and Domain Structure of and Irradiation Effect on Ferroelectric Glycine Sulfate Single Crystals."
(10) W. J. Merz; *J. Appl. Phys.* **27**, (1956) 938. "Switching Time in Ferroelectric $BaTiO_3$ and Its Dependence on Crystal Thickness."

* * * *

REVIEW OF THE ELECTRICAL COMMUNICATION LABORATORY
VOLUME 9, NUMBERS 5-6, MAY-JUNE, 1961

U.D.C. 541. 127 : 546. 36′86 : 537. 312. 5

Formation of Cesium Antimonide. II. Formation of Cs_3Sb at Temperatures below 130°C*

Kiyoshi MIYAKE†

Films of cesium antimonides were prepared in the temperature range from 28° C to 130° C under a constant vapor pressure of cesium. The changes in electrical properties of the samples produced by the reaction of cesium with antimony were observed in each stage of the formation process. Electrical resistivity vs. reaction time curves of the samples prepared at all temperatures had two maxima and one minimum, whereas changes in photoresponse depended on the preparation temperature; photoresponse vs. reaction time curves for samples prepared at temperatures above 100° C had a maximum near the second maximum of the resistance, and the curves for samples prepared at temperature below 85° C had two maxima and one minimum which occurred near the corresponding points on the resistance curves. The material at the second maximum on the resistance curves was supposed to be Cs_3Sb*, whose maximum resistivity at 0° C and maximum photoresponse at room temperature were 1530 ohm-cm and 55 μA/lumen, respectively. The maxima and the minimum of the resistance vs. time curves are discussed in connection with the electrical resistivity of films of the cesium-antimony system.*

1. Introduction

Cesium antimonide used as the cathode material of photo-tubes has been studied by many investigators, but there are some discrepancies in the reported values of electrical conductivity, thermal activation energy associated with conductivity, and photoelectric response of this material; which would arise from variation in the composition of the samples used. Usually, cesium antimony photocathodes are prepared at about 150° C by reacting cesium with an evaporated antimony film until the photoresponse attains a maximum value. Investigators, however, have not always prepared the cathode at the same temperature.

Prilezhaev(1) prepared the cathode in the temperature range from 140° C to 260° C; Sommer,(2) from 140° C to 190° C; Janes and Glover,(3) and Janes,(4) from 150° C to 170° C; Khorosh,(5) from 120° C to 180° C; Fainshtein,(6) at 210° C; Suhrmann and Kreisen,(7) at 160° C; and Wallis,(8) from 120° C to 150° C. Suhrmann and Kreisen, and Menshikov,(9) reported that an intermediate compound was formed by treating an antimony film with cesium vapor at room temperature. Janes and Glover(3) also observed that an alloy of antimony and cesium was formed at room temperature.

In general, at the usual preparation temperatures, the reaction rate of cesium with the antimony film is very fast, so that it is difficult to observe the course of the reaction in detail. Therefore, information on the reaction of cesium with antimony and on the characteristics of the compounds with definite composition which are formed in the course

---

* MS received by the Electrical Communication Laboratory on 17 January 1961. Originally published in the *Review of the Electrical Communication Laboratory*, Vol. 9, Nos. 5-6, pp. 287–294, 1961.

† Semiconductor Research Section.

of reaction will be obtained by measurements of the resistance and photoresponse of films only if cesium antimonide is formed at a sufficiently slow reaction rate at lower tempratures.

The purpose of this investigation is to study the following matters: (1) Preparation of cesium antimonide having nearly stoichiometric composition at a slow reaction rate under a constant vapor pressure of cesium. (2) Possibility of the formation of cesium antimonide at temperatures ranging from 130°C to room temperature. (3) Electrical resistivity and its temperature dependence, and photoelectric response of the films at definite stages in the course of formation. (4) Activation energy of the reaction required for the formation of compounds.

2. Experimental Procedures

Fig. 1 shows the reaction tube used. Cesium antimonide was prepared on a glass plate (10 mm × 8 mm) mounted in a cylindrical borosilicate glass envelope (4 cm diam. and 12 cm long.). The glass plate has platinum foil leads imbedded into opposite edges which were platinized over to obtain a good contact with the film to be deposited. The purity of the antimony used was higher than 99.99%.⁽¹⁰⁾ Metallic cesium was made by reducing cesium chromate with silicon.

Experimental tubes were prepared by the following procedures: (i) The whole tube was evacuated and baked at a temperature of 300° C to 350° C for 3 hours. (ii) Antimnoy was evaporated onto the glass plate from tungsten coils. The thickness of the film evaporated on the glass plate (900 A) was calculated from that of a film evaported on the inner wall of the envelope, which was determined by comparing it with the known thickness of a reference film. (iii) Metallic cesium was distilled through a condensing chamber, and stored in a side tube (1.5 cm diam. and 12 cm long.). (iv) The assembly including the side tube was tipped-off from the evacuating system at a vacuum better than 1×10^{-7} mm Hg. The maximum deviation in thickness of the films evaporated on the glass substrate was estimated to be less than 50 A. Antimony was evaporated at a sufficiently slow rate to obtain a uniform film thickness.⁽¹¹⁾

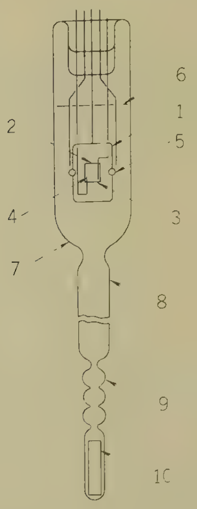


Fig. 1—Reaction tube used for the preparation of cesium antimonide.

(1) Anode. (2) Glass plate on which the sample is prepared. (3) Platinized area. (4) Platinum foil lead burned in glass. (5) Evaporator. (6) Mica support. (7) Hard glass envelope. (8) Side tube in which metallic cesium is stored. (9) Condensing chamber. (10) Nickel pellet containing cesium chromate and silicon.

In general, cesium vapor makes contact with an antimony film, followed by the formation of a nonmetallic compound between them. The formation rate of the compound would be determined by either the diffusion rate of the cesium atoms through the film or the rate of arrival of the cesium atoms on the surface of the film. If the rate of arrival is larger than the diffusion rate, the formation rate of the compound will be determined by the diffusion rate of the cesium atoms. In the present work, the whole tube including the side tube containing a large amount of metallic cesium was heated at a constant temperature, and the reaction of cesium vapor with the antimony film took place under the conditions mentioned above. Consequently, the cesium content in the film is not proportional to the reaction time. The area of the inner surface of the side tube and the diameter of the neck at its upper end were determined considering the rate of supply of cesium vapor.

The experimental procedures for the formation of cesium antimonide were as follows: The assembly was heated in an electric furnace kept at a given temperature. With both ends of the electric furnace closed, the maximum deviation in the temperature distribution along the whole length of the assembly could be held below 2° C. The temperature of the furnace was automatically controlled and was kept at a temperature constant to within ±0.5° C.

After heating for a given time, the assembly was taken out of the furnace, and immediately the side tube was cooled with dry ice or liquid oxygen in order to distill any nonreacted cesium into it; thereafter the electrical resistance and the photoresponse of the sample were measured. Such experimental procedures were repeated until the sample had the usual photoresponse of cesium-antimony photocathodes and had a high electrical resistance (as shown in Figure 2).

The electrical resistance of the samples was measured at 0° C in total darkness, and the photoresponse to the ligth of a tungsten lamp (2847° K) was measured at room temperautre. The temperature dependence of the resistance was measured after the samples had been stored in total darkness for 16 hours. During the measurement of the electrical properties of the samples, the side tube was cooled with dry ice to prevent the evaporation of metallic cesium.

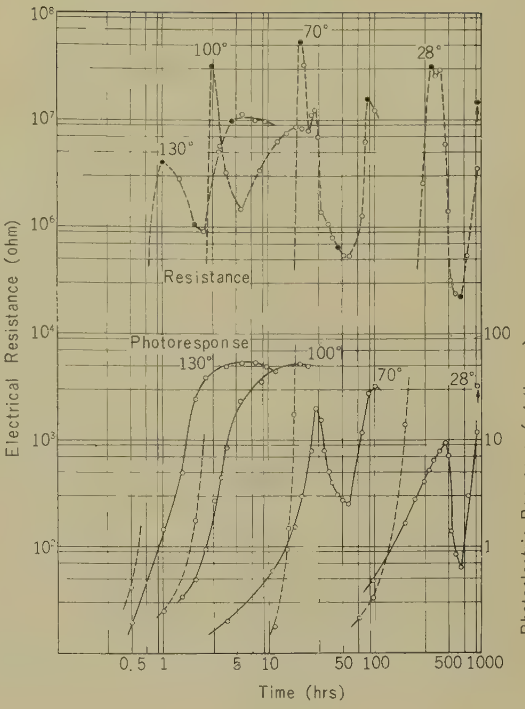


Fig. 2—Changes in the electrical resistance and photoresponse of the samples during the formation of Cs_3Sb. Resistance was measured at 0°C, and photoresponse was determined for the light of a tungsten lamp (color temperature 2847° K).

3. Experimental Results

Figure 2 shows typical changes in the electrical resistance and photoresponse of the sample with successive heating at given temprtures. At any heating temperature, similiar behavior of resistance change was observed, i.e. the resistance vs. time curves had two maxima and one minimum. On the other hand, the change in photoresponse depended on the heating temperatere; above 100° C the response increased monotonically with increasing time and reached its usual value (about 40 μA/lumen) before the resistance reached the second maximum of the curve, which agrees with the results obtained by Miyazawa,(12) below 85° C it changed together with the resistance. It should be noted that even at room temperature a material having a usual value of photoresponse of the antimony-cesium photocathodes was formed after a long time.

Such relations between the electrical properties and the reaction time (heating time) were reproducible at any reaction temperature, except for the reaction at 130° C. In the experiment at 130° C four samples were used, two of which did not have the large first maximum of resistance. The first maximum appearing

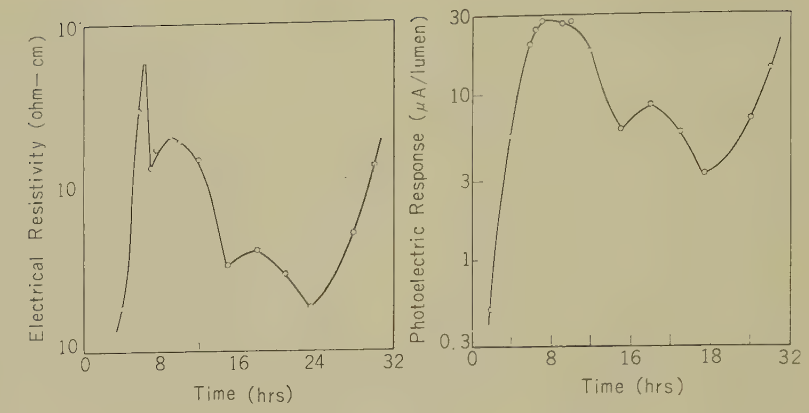


Fig. 3—Changes in the electrical resistance and photoresponse between the first maximum consisting of two small peaks and the minimum of the resistance (reaction temperatare 85° C).

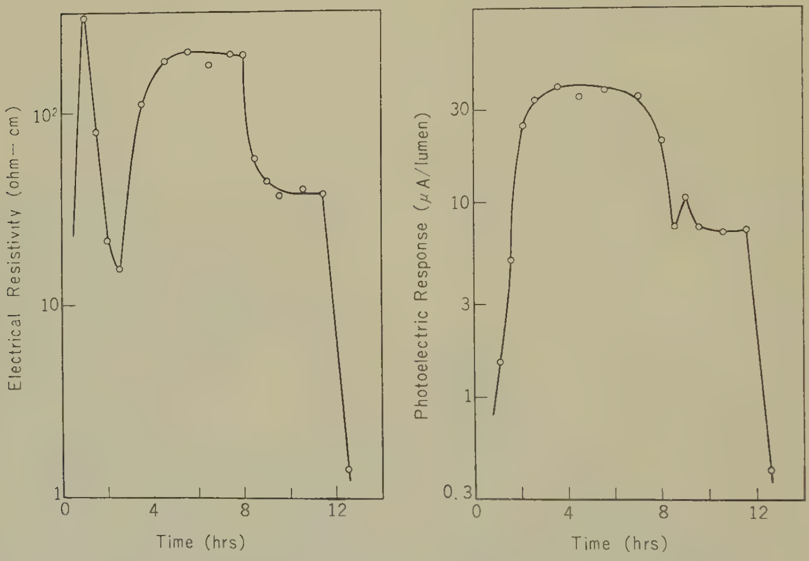


Fig. 4—Changes in the electrical resistance and photoresponse during prolonged heating at 130° C.

in the reactions below 85° C had a structure consisting of two small peaks as shown in Figure 3.

The minimum of the resistance vs. time curves occurred for all reaction temperatures. However, the resistance of the samples prepared at lower reaction temperatures did not change monotonically from the first maximum to the minimum, but a small hump appeared between them.

The second maximum of the resistance and the usual photoresponse of the antimony-cesium photocathodes were obtained only when the nonreacted cesium vapor was completely distilled in the side tube cooled by liquid oxygen; otherwise the resistance and the photoresponse were considerably reduced. When the experimental procedures mentioned previously were further repeated beyond the second maximum of the resistance, the resistance and the photoresponse decreased with increasing time as shown in Figure 4. Generally, when the resistance of the samples approached the second maximum, their photoresponses were activated by drawing large photocurrents from them. At the same time, their resistances also increased. These increases in the resistance and the photoresponse were shown by arrows in Figure 2.

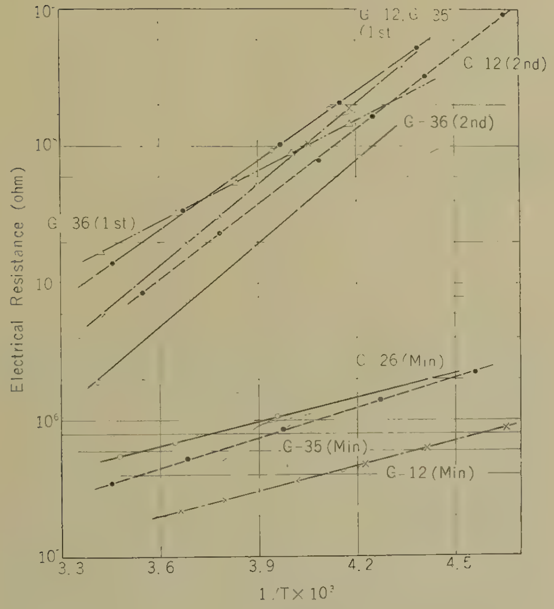


Fig. 5—Electrical resistance of the samples as a function of the reciprocal of the absolute temperature at the first and second maximum and the minimum of the resistance vs. time curves. (The stages at which the temperature dependence of electrical resistance was measured are shown in Fig. 2 by filled-in circles on the curves.)

Figure 5 shows typically the relation between the logarithm of electrical resistance and the reciprocal of the absolute temperature obtained at three stages of the films, namely the first maximum, the minimum, and the second maximum of the resistance vs. time curves. Filled-in circles on each curve in Figure 2 show the three stages at which the temperature dependence of the resistance was measured.

A drift phenomenon in the electrical resistance of the compounds having a compsition near the second maximum was observed. In such cases the measurement was carried out, with the applied voltage between the two leads to the film reduced to less than 0.2 V from the usual value 1.5 V, after the tube had been stored in total darkness for more than 16 hours.(13)~(15)

Electrical resistivity, thermal activation energy associated with conductivity, and photoresponse of the films at the two maxima mentioned previously are tabulated in Table 1. The electrical resistivities were caluclated using the thickness of evaporated antimony films (900 A). The electrical resistivity and the thermal activation energy at the first maximum tabulated in Table 1 are those which were measured at the first peak in the maximum, except for those of the samples prepared at temperatures above 100° C.

4. Discussion

In our previous work on the electrical resistivity of the films of the cesium-antimony system,(10) the heating of the films was continued at a constant temperature until no change in electrical resistance was produced by further heating. It was then assumed that the composition of the films was homogeneous. On the other hand, in the present work, a small concentration gradient of cesium atoms would be present in the film at stages in the course of the reaction. This is because the heating of the samples, as well as the supply of cesium vapor, was stopped and the samples were immediately cooled to room temperature for every measurement of the electrical properties.

It has been generally accepted that the composition of a cesium antimonide having the highest photoresponse corresponds to the formula Cs_3Sb. In the present work, it was found that the highest photoresponse occurred at the second maximum of the resistance vs. time curves as shown in Fig. 2 and Fig. 4. Moreover, the electrical resistivity and the

Table 1

Activation Energy, Resistivity, and Photoresponse
Second Maximum and at the Minimum

Sample number	Preparation temperature (°C)	First maximum		
		Activation energy (ev)	Resistivity at 0° C (ohm-cm)	Photo-response (μA/lumen)
G-12	28	0. 62	740	9. 6
G-33	39	0. 60	720	14
G-31	50	0. 60	900	17
G-26	70	0. 62	1300	20
G-35	85	0. 62	760	27
G-36	100	0. 62	740	—
G-37	130	0. 52	90	—
G-19	130	—	380	—

thermal activation energy associated with conductivity at the maximum tabulated in Table 1 agree with those of $Cs_{3.03}Sb$ obtained previously. Thus it is supposed that the second maximum of the curves arises from the formation of Cs_3Sb.

As shown in Fig. 4, the resistivity and the photoresponse of the samples are generally decreased by further heating beyond the second maximum of resistance. It was reported previously that compounds containing an amount of cesium larger than Cs_3Sb could be formed. Also it has been found that the partial pressure of cesium over Cs_3Sb at 130° C measured by Langmuir's positive ion method is about 3×10^{-7} mmHg.(16),(17) In the present work, the samples were prepared under the saturated vapor pressure of cesium (4×10^{-3} mmHg at 130° C). Thus the decreases in the electrical resistivity and the photoresponse, which occur after the resistance reaches the second maximum, would be caused by the formation of the compound having an amount of cesium larger than Cs_3Sb.

It was suggested previously that there were compounds in the cesium-antimony system: $CsSb$, Cs_3Sb_2, Cs_2Sb, and Cs_3Sb; and it was expected that the peaks of the resistance corresponding to the formation of the compounds would appear in the resistance vs. time curves. However, no peaks appeared between the first and the second maximum of the curves, except that in the reaction below 85° C the first maximum of resistance splits into two peaks as shown in Fig. 3.

Therefore another experiment was carried out at room temperature in order to observe the change in the resistance near the first maximum in detail using thin antimony films (300 A). It was found that the first maximum of the curve for the films consisted of two small peaks. The electrical resistivities and activation energies tabulated in the third column of Table 1 agree with those of $Cs_{1.02}Sb$ reported previously. Thus, from the results obtained by several investigators(2),(8),(12) and the facts mentioned above, the compounds at the first peak and the second peak on the maximum are deduced to be $CsSb$ and Cs_3Sb_2, respectively.

On the other hand, it was difficult to observe the peak on the resistance vs. time curves corresponding to the formation of Cs_2Sb. This is probably due to the following condition: In general, a small concentration gradient of cesium atoms will be present in

OF THE SAMPLES AT THE FIRST MAXIMUM AND THE
ON THE RESISTANCE VS. TIME CURVES

Minimum			Second maximum		
Activation energy (ev)	Resistivity at 0°C (ohm-cm)	Photoresponse (μA/lumen)	Activation energy (ev)	Resistivity at 0°C (ohm-cm)	Photoresponse (μA/lumen)
0. 24	5. 2	0. 65	0. 71	340	32
0. 21	7. 2	3. 3	0. 71	1530	33
0. 23	5. 4	4. 9	0. 71	500	42
0. 24	15	2. 5	0. 73	360	33
0. 28	11	5. 8	0. 77	410	54
——	34	24	0. 78	150	51
0. 50	21	40	0. 78	210	55
——	21	35	——	270	40

the film. Consequently, it is probable that a layer of $Cs_{2\cdot02}Sb$ with high resistivity, 2.85×10^3 ohm-cm, and a thin layer of $Cs_{2\cdot12}Sb$ with low resistivity, 7.99×10^{-2} ohm-cm, are simultaneously present in the film. If a layer of $Cs_{2\cdot12}Sb$ whose thickness is larger than 10 A is present in the film, the resistance will be reduced to the order of 10 ohms even though all other layers except the $Cs_{2\cdot12}Sb$ are insulators. It is therefore supposed that the peak due to the formation of Cs_2Sb disappears and the minimum due to the formation of the $Cs_{2\cdot12}Sb$ layer appears. The activation energies of the films at the minimum in the formation below 85° C are in good agreement, although the resistivity increased with increasing temperature of formation (Table 1).

From a plot of the logarithm of the reciprocal of the time required for attaining the first peak on the first maximum and one of the time required for reaching from the peak to the second maximum to the reciprocal of the absolute temperature of the reaction, linear relations were obtained as shown in Fig. 6 and are given by the following equations.

$$1/t = 7.24 \times 10^7 \exp(-14300/RT),$$

for the first peak,

$$1/t = 3.39 \times 10^5 \exp(-11500/RT),$$

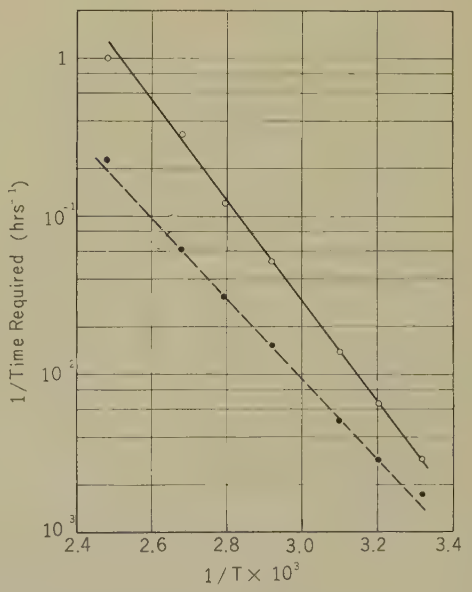


Fig. 6—The time required for attaining the first maximum and the time between the first and the second maximum (filled-in circles) as a function of reaction temperature (°K).

for the second maximum, where t is the time in hours.

As shown in Fig. 2, the change in the photoresponse depends on the temperature at which the samples are prepared. This temperature dependence is caused by the difference in the concentration gradient of cesium atoms in the films, because the concentration gradient is determined by the relative magnitude of the diffusion rate of cesium atoms to the supply rate of cesium atoms.

Recently, Sommer reported that the dark resistivity of Cs_3Sb measured at 150° C was still increasing at the point of peak photoresponse and reached a maximum only if further cesium was added.(18) Similar behavior in the dark resistance observed at 170° C was reported by Imamura.(19) However in the present work, the dark resistance of the samples measured at 0° C reached a maximum before attaining a maximum photoresponse, which was observed near the first maximum of the resistance as well as near the second maximum. This disagreement in the behavior is mainly due to the difference in the temperature at which the resistance was measured, in the thickness of the films used, and in the method of preparing the samples. The thermal activation energy associated with conductivity of the cesium antimonides varies with different compositions. It is therefore probable that the resistance of a Cs_3Sb film with a small amount of excess cesium is higher than that of Cs_3Sb, which depend on the temperature at which the resistance is measured.

Acknowledgement

The author wishes to express his hearty thanks to Mr. M. Sugihara for helpful discussions in preparing this paper, and to Dr. T. Niimi, the Chief of the Semicondnctor Research Section, for encouragement.

References

(1) S. Prilezhav: Antimony-Cesium Photocathode, *Zhur. Tekn. Fiz.* **9.** 1439 (1939); C. A. **34,** p. 1909, 1940.

(2) A. Sommer: Photoelctrically Sensitive Layers, *U. S. Patent* 2192, 418 Mar. 5, 1940.

(3) R. Janes and A. Glover: Method of Manufacturing Photoelectric Tubes, *U. S. Patent* 2401, 735 June 11, 1946.

(4) R. Janes: Photoelectric Tubes, *U. S. Patent* 2431, 401 Nov. 1947.

(5) D. Khorosh: Photoelectric Properties of Cs-Sb Photocathodes in Transmitted Light as a Function of Thickness, *Zhur. Tekn. Fiz.* **17,** 341 (1947); C A. **42,** p. 816, 1948.

(6) S. Fainshtein :Photoelectric Amplifier with Sb-Cs Cathode, *Zhur. Tekn. Fiz.* **18,** 39 (1947); C. A. **45,** p. 10066, 1951.

(7) R. Suhrmann and G. Kreisen: Uber die Emissionzentren der an Intermetallischen Verbindungen Ausgelosten Photoelektronen, *Z. Elektrochem.* **54,** Heft 5, s. 349, 1950.

(8) G. Wallis: Optiche und Elektriche Messungen an Caesium-Antiomon Schichten Vershieden Zusammensetzung, *Ann. Phsik* **17,** Heft 6, s. 401, 1956.

(9) M. Menshikov: Photoelectric and Optical Properties of Cs-Sb Films of Varying Composition, *Zhur. Tekn. Fiz.* **17,** 579 (1947); C. A. **42,** p. 1118, 1948.

(10) K. Miyake: Electrical Resistivity of the Films of Cs-Sb System, *J. Appl. Phys.* **31,** No. 1, p. 76, 1960.

(11) See reference (3). Evaporation rate in the present work is lower than that adopted by Janes and Glover.

(12) H. Miyazawa *et al.*: On the Composition of Antimony-Cesium Photocathodes, *J. Phys. Soc. Japan*, **7,** No. 6, p. 647, 1952'

(13) See reference (10). We observed a drift phenomenon in the electrical resistance of the samples whose compositions were $Cs_{2.67}Sb$, $Cs_{2.75}Sb$, and $Cs_{3.30}Sb$.

(14) W. Widmaier and R. Engstrom: Variation of the Conductivity of the Semitransparent Cesium-Antimony Photocathodes. *R C A Rev.* **16,** No. 1, p. 109, 1955.

(15) W. Harper and W. Choyke: The Resistance of Semitransparent Photocathodes, *J. Appl. Phys.* **27,** No. 11, p. 1358, 1957.

(16) J. Taylor and I. Langmuir: Vapor Pressure of Cesium by the Positive Ion Method, *Phys. Rev.* **51,** No. 6, p. 753, 1939.

(17) To be published in *J. Appl. Phys.*, **32,** 6, 1961.

(18) A. Sommer: n-Type and p-Type Conduction in Alkali-Atimonide Photoemitters, *J. Appl. Phys.* **29,** No. 11, p. 1568, 1958.

(19) S. Imamura: Electrical Properties of Alkali Antimonides, *J. Phys. Soc. Japan,* **14,** No 11, p. 1491, 1959.

REVIEW OF THE ELECTRICAL COMMUNICATION LABORATORY
VOLUME 9, NUMBERS 5-6, MAY-JUNE, 1961

U.D.C. 539. 23 : 546. 36′86 : 537. 312. 5

Formation of Cesium Antimonide. III. Logarithmic Growth of Cesium Antimonide Films on Antimony*

Kiyoshi MIYAKE†

Surface reaction of vacuum evaporated antimony films with cesium vapor was studied at room temperature in relation to the formation of cesium antimonides. Film in the thickness range between 250 A and 1100 A were used. Changes in the electrical resistance of these films produced by the reaction were observed, and the thickness of the growing layer of compound was determined using the thickness of the non-reacted antimony layer calculated from its electrical resistance during the course of the reaction. It was then found that the following relation between the thickness of the growing layer, x, and reaction time, t, held $t \gg c$. $x = k \log(t/c+1)$, where k and c are constants.

1. Introduction

Many metals react with gases such as oxygen, the halogens, and sulfur at high temperature, and some even at room temperature, followed by the formation of the oxide or other compound on their surfaces. In the investigation of the surface reactions of metals; several methods, such as optical, weighing, and electrical methods, have been used to estimate the thickness of the surface layer formed, and to deduce the end of the reaction.(1)

The reaction of antimony films with cesium vapor is also one of the surface reactions of metals, and it has been suggested in our previous work(2) that the reaction at the early stage of formation of cesium antimonide films is a kind of reaction in which an electropositive gas; namely, cesium vapor makes contact with the antimony film which behaves as an electronegative metal, followed by the formation of a nonmetallic compound between them. The compound has a very high resistivity, and has been suggested to be CsSb.(3),(4),(5)

In our previous work, it was found that a film of $Cs_{0\cdot91}Sb$ has a very low resistivity, about 10^{-3} ohm-cm, and showed a metallic character; whereas a film of $Cs_{1\cdot02}Sb$ had a high resistivity, about 10^3 ohm-cm, and exhibited the characteristics of a semiconductor. And it was suggested that the film of $Cs_{0\cdot91}\cdot Sb$ consisted of a thick CsSb layer and a thin layer of non-reacted antimony or dilute cesium-antimony alloy.

Therefore, if the increase in electrical resistance of an antimony film at the early stage of the formation of cesium antimonide, which is produced by the reaction of cesium with the antimony, is due to decrease in thickness of the antimony layer, a relation between the thickness of growing layer of the compound and the reaction time will be obtained. In view of the above considerations, this work was carried out.

---

* Ms in Japanese received by the Electrical Communication Laboratory on January 10, 1961. Originally published in the *Kenkyû Zituyôka Hôkoku* (*Electrical Communication Laboratory Technical Journal*), Vol. 10, No. 6, pp. 1233–1236, 1961.

† Semiconductor Research Section.

2. Experimental Procedures and Results

The experimental reaction tube used consists of an evaporator and a square glass plate mounted in a cylindrical glass envelope (4 cm diam. and 10 cm long). A side tube consisting of a U-shaped trap for cesium were connected to one end of the glass envelope. A weighed fragment of antimony was located in the evaporator, from which it was evaporated on the glass plate. The trap for cesium vapor was cooled with liquid nitrogen or oxygen when the reaction was to be stopped. The tubes were evacuated and baked out under the same conditions as those mentioned in our previous report. The evaporation rate of the antimony ranged from 0.01 to 0.002 mg per cm^2 per min to obtain a film of uniform thickness.(7) The antimony was evaporated at room temperature in a vacuum better than 3×10^{-8} mm Hg. The thickness of the antimony films used was calculated using a known amount of antimony, and the distance between the evaporator and the substrate was varied from 250 A to 1100 A in order to study crystallized films.(7),(8) The dimensions of the film used were 4 mm long and 10 mm wide.

The experiments were carried out under conditions in which the rate of supply of the the cesium vapor was higher than the formation rate of the compound, in order that the latter was not determined by the former. Consequently the leakage current due to adsorbed cesium on the glass surfaces, which appears in the measurement of the electrical resistance, increases with increasing time of reaction. The leakage current, therefore, was monitored by measuring the resistance between two leads connected to the sample and the evaporator, and the resistance was estimated to be about 10^9 ohms at the final stage of the reaction.

The original resistance of the antimony films was calculated using their dimensions and the electrical resistivity in bulk.* Prior to the experiment, the total resistance between two leads attaches to the opposite edges of the glass plate was measured order to determine any resistance other than the original resistance of the antimony film to be studied. The resistances except for the original resistance were determined by subtracting the original resistance from the total resistance measured, and were found to be less than 30 ohms.

The reaction of cesium vapor with the antimony films was carried out at room temperature in total darkness in order to eliminate the effect of light on the reaction rate. The electrical resistance of the samples was measured under the same conditions. All the measurements of resistance were carried out using a potentiometer and a standard resistance.

The resistivity of the compound layer grown on the surface of antimony is about 10^7 times that of antimony in bulk, so that the resistance of the samples measured at the early stage of the reaction may be regarded as that of the non-reacted antimony film. Then, neglecting an increase in volume of the film which may be produced by the reaction, the thickness of the growing layer was determined by subtracting the thickness of the non-reacted antimony layer calculated from the resistance of the sample in the course of reaction from the original thickness of the antimony film.†

Fig. 1 shows schematically the sample at the initial and intermediate stages of the reaction. Fig. 2 shows the typical relations between the electrical resistance of the samples and the reaction time. The resistance of the

---

* Leverton and Decker have studied the resistivity of thin films ot antimony prepared by distillation, and reported that the resistivity of the annealed films was 1.28 times that of bulk antimony. (*Phys. Rev.* **80**, 732 (1950)).

† The thickness of the growing layer should be proportional to thickness of the antimony layer consumed by the reaction, provided that a homogeneous layer of the compound is formed. The thickness determined by the subtraction is that of the consumed antimony layer, so that the correct value of the thickness of the growing layer will be obtained by multiplying it by a factor. We, however, have no quantitative information on the volume change in the formation of CsSb from the components. Therefore, we adopted this method, by which the linear relation between x and $\log t$ shown in Fig. 2 is not altered.

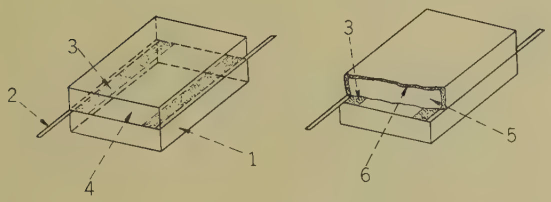


Fig. 1—Schematic diagram of the film at the initial and intermediate stages of the reaction.

(1) Glass substrate. (2) Platinum foil lead. (3) Platinized area for obtaining a good contact between the platinum foil and the evaporated antimony film. (4) Evaporated antimony film before reaction. (5) Non-reacted antimony layer. (6) Surface layer of cesium antimonide grown on antimony.

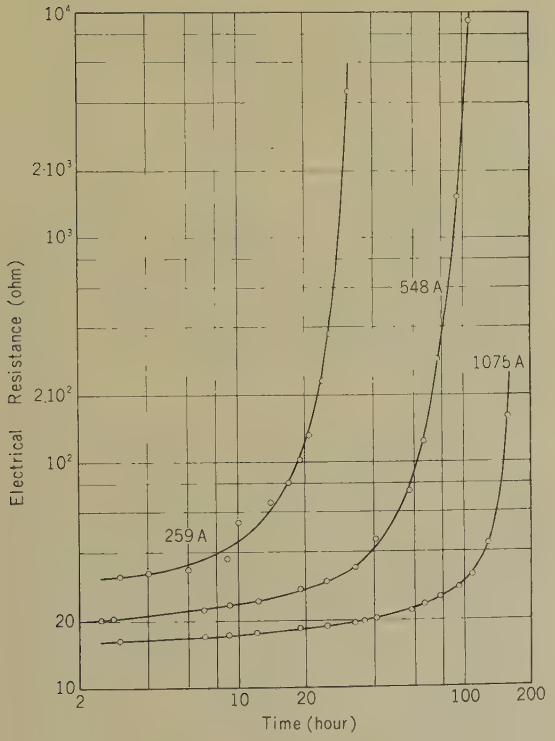


Fig. 2—Changes in the electrical resistance of samples produced by the reaction of cesium with antimony. Numbers on each curve show the thickness of the original antimony films.

samples increases with increasing time and reaches a maximum at which the resistance is above 10^7 ohms.

Fig. 3 shows the thickness of the compound layer grown on antimony films as a function of the logarithm of reaction time. As shown in the figure, the thickness of the growing layer x in A and time t in hours are related by the following equation until the thickness of non-reacted antimony layer is reduced to about fifty Angstroms.

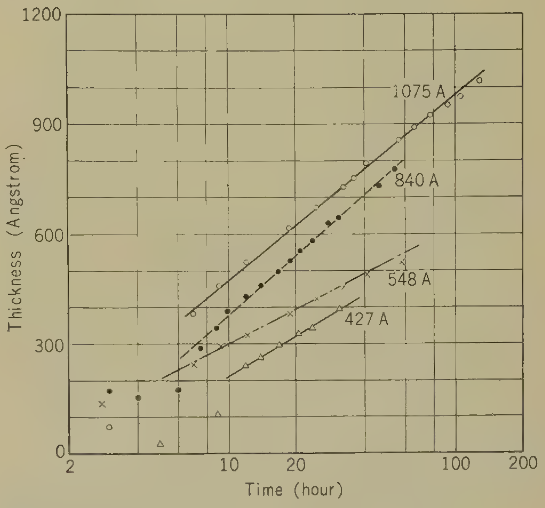


Fig. 3—Relations between the thickness of the compound layer grown on antimony film and the reaction time at room temperature. Numbers on each curve show the thickness of the original antimony films.

$$x = k \log\left(\frac{t}{c} + 1\right), \qquad (1)$$

where k and c are constants.

Generally there is an induction period of the reaction, and for $t \gg c$, the functional relation between x and t shown by Eq. (1) is reproducible. However, k and c were not always constants; with different samples k varied from 310 A to 520 A, and c varied from 1.1 hrs to 3.2 hrs. With thin films below 300 A, and for the induction period the plots between x and $\log t$ were scattered.

3. Discussion

As shown in Fig. 3, the linear relation between x and $\log t$ holds for $t \gg c$. However, after the thickness of the non-reacted antimony layer had been reduced to below 50 Å, measured points departed from the curves given by Eq. (1), and a comparatively long time was required until the whole layer of the sample had completely changed to the compound. Therefore, the formation mechanism of the compound at this stage may be different from that at the early stage.

At the early stage of the reaction, a remarkable induction period was generally observed. For this stage, the resistance of the films was only a small part of the total resistance, and also only a small increase in resistance was produced by the progress of the reaction. Consequently, a comparatively large error in the calculated thickness of the growing layer resulted and it was estimated to be about 15 A for a layer with a thickness of 300 A. The error increases with decreasing thickness of the layer. Therefore, more precise measurements of the electrical resistance are required for obtaining an exact relation between x and t at this stage.

Many investigators have studied evaporated antimony films, and found that the structure of the films and their electrical properties were dependent on the evaporation rate of the metal, the thickness of films, and the temperature of the substrate.(9) Frimer(10) investigated vacuum-evaporated thin layers of antimony and layers of cesium on antimony by means of the electron microscope and electron diffraction, and found that the slowly evaporated antimony layer apparently became continuous at a thickness of 300 A. He also found that the condensation of the first layer of cesium on the antimony film changed its appearance from homogeneous granular to heterogeneous. His results suggest that our samples also had irregularities in microscopic surface conditions and variations in the granular structure of the film. Therefore, the variations of rate constant and induction period with different samples may be related to these matters.

Pilling and Bedworth(11) have pointed out that in the surface reaction of metals with gases the volume of the compound formed must be larger than that of the reacting metal in order to form a protective surface layer free from pores. Unfortunately, we have no information on the volume change of CsSb during the formation from its components. However, Jack and Wachtel(12) have investigated the crystal structure of cesium antimonide with various compositions by X-ray crystal analysis, and observed that samples of the powder had increased in volume when cesium vapor reacted with antimony. Munesue (13) has investigated the change in the thickness of films during the formation of cesium antimonide by an optical method and found that the thickness of the film increased to several times that of the original antimony film. From their results the surface layer of cesium antimonide formed on the antimony films is reasonably assumed to be compact.

If the rate determining step of the formation reaction of the compound is a diffusion process of cesium or antimony atoms through the layer, and the rate of growth is proportional to a concentration gradient in the layer, a parabollic equation for the growth of the compound will be obtained. We, however, obtained a logarithmic equation for the growth. Therefore, the formation rate of the compound may be determined by other processes.

Acknowledgment

The author wishes to express his hearty thanks to Dr. T. Niimi, the Chief of the Semiconductor Research Section of the Electrical Communication Laboratory, for encouragement.

References

(1) W. Jost: Diffusion in Solids, Liquids, Gases, p. 340, 1952, *Academic Press Inc.*

(2) K. Miyake: Formation of Cesium Antimonides. I. Electrical Resistivity of the Film of Casium-Antimony System, *J. Appl. Phys.* **31**, No. 3, p. 76, 1960.

(3) A. Sommer: Photoelectric Alloys of Alkalic Metals, *Proc. Phys. Soc.* (*London*), **55,** No. 2, p. 145, 1943.

(4) H. Miyazawa et al.: On the Composition of Antimony Cesium Photocathodes, *J. Phys. Soc. Japan*, **7,** No. 6, p. 647, 1952.

(5) G. Wallis: Optische und Electrische Messungen an Caesium-Antimon Schichten, *Ann. Physik*, **17,** Heft 6, s. 401, 1956.

(6) R. Janes and A. Glover: Method of Manufacturing Phototubes, *U. S. Patent* 2,401,735 June 11, 1946.

(7) A. Gotzberger: Uber die Kristallisation auf Aufgedampter Antimonschichten, *Z. Physik*, **142,** Heft 2, s. 182, 1955.

(8) W. Latmer: Crystallization of thin Sb Films, C.A. **40,** No. 1, p. 1371, 1946; *Hele. Phys. Acta*, **18,** 232, 1945.

(9) J. Choen: Semiconducting Films of Antimony, *J. Appl: Phys.* **25,** No. 6, p. 798, 1954.

(10) A. Frimer: Electron Microscopic Investigation of Sb-Cs layers, C.A. **43,** No. 2, p. 2481, 1949; *Doklady Akad. Nauk. S.S.SR.*, **63,** 255, 1948.

(11) N. Pilling and R. Bedworth: *J. Inst. Metals*, **29,** 539, 1923. See reference (1).

(12) K. Jack and M. Wachtel: The Characterization and Crystal Structure of Cesium Antimonide, a Photoelectric Surface Material, *Proc. Roy. Soc.* (*London*), **A 239,** No. 1216, p. 46, 1957.

(13) S. Munesue: Private Communication.

* * * *

Review of the Electrical Communication Laboratory
Volume 9, Numbers 5-6, May-June, 1961

U.D.C. 534. 43. 062 : 546. 71'72
669. 15[543. 062 : 546. 7

A Method of Quantifying Fe^{2+} and Mn^{3+} in Sintered Mn-Zn Ferrites*

***Hideo NAGATO*†**

In order to determine the quantity of Mn^{3+}, *"experimental designs" were applied in inspecting the analytical conditions using* H_3PO_4–$HCLO_2$ *series solution. The ferrite samples were dissolved in a mixed solution of poly phosphoric and perchloric acid at high temperature (265° C) in a short time (10 sec.), and the extinction of* Mn^{3+} *in the solution was measured. To determine the quantity of* Fe^{2+}, *the samples were dissolved in poly phosphoric acid solution containing* Mn^{3+} *at the same temperature, and the increase or decrease in the extinction of* Mn^{3+} *was measured. The quantity of* Fe^{2+} *was found from this extinction change and the above mentioned results. The method was applied to the determining of the composition of sintered iron oxides, sintered manganese oxides, and sintered Mn-Zn ferrites having different compositions prepared by different sintering and quenching conditions.*

Introduction

Sintered materials of the MnO-ZnO-Fe_2O_3 ternary oxide series are widely used for carrier frequency telephone apparatus. The magnetic characteristics of these sintered oxides vary according to the sintering conditions, and the reason for this has been generally thought to be due to the fact that the valencies of iron oxide and manganese oxide easily vary to higher than divalent, and the amounts and the distributions of Fe^{2+}, Fe^{3+}, Mn^{2+}, Mn^{3+} in the oxides change with oxidation or reduction due to the sintering treatment.

Methods of quantitative analysis of metal ions which have this spinel form with various valencies has not been given in domestic and foreign references. The author has studied a chemical quantitative analysis method of the valency for control of the magnetic characters of sintered materials. Next, he made sintered iron oxides, sintered manganese oxides, and sintered Mn-Zn ferrites under various conditions, after which he performed quantitative analysis and deduced the compositions of each sintered material from the analytical results.

1. Analysis of Mn-Zn Ferrites

1.1. Absorption Curve of Mn^{3+} and Calibration Curve

Although many methods of preparation of Mn^{3+} standard solution have been reported, the author found a new method which consists of the decomposition of $KMnO_4$ with a heated mixed acid solution ($H_3PO_4/HClO_4=30/1$ by vol.). Mn^{3+} was also prepared from manganese ϕ oxide and manganese carbonate as well as from $KMnO_4$.

This method is a new method for preparing Mn^{3+} standard solution and can be briefly explaining as follows: 30 ml. of the mixed

---

* MS received by the Electrical Communication Laboratory on October 15, 1960. Originally published in the *Kenkyû Zituyôka Hôkoku* (*Electrical Communication Laboratory Technical Journal*), N.T.T., Vol. 10, No. 4 (II), pp. 629-656, April 1961.

† Semiconductor Research Section.

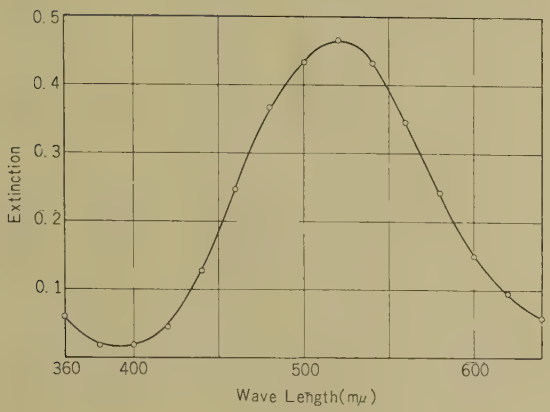


Fig. 1—Extinction-core entration curve of Mn^{3+}.

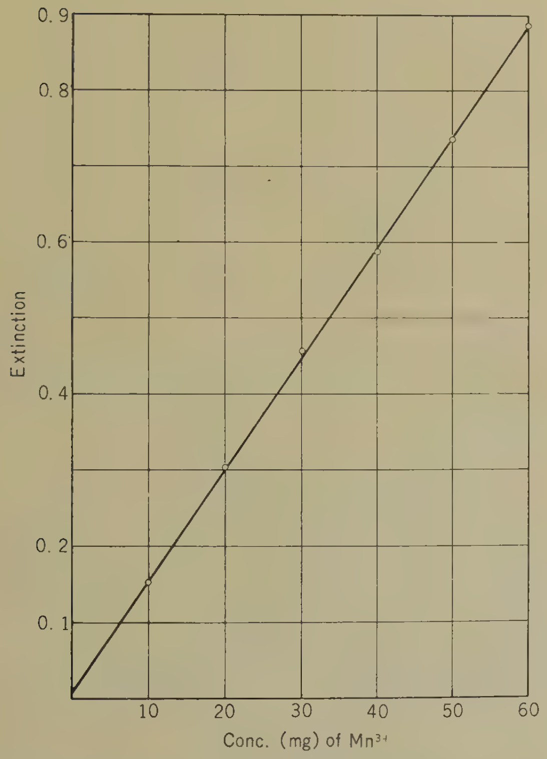


Fig. 2—The relation curve between the oxidation temperature of Mn^{2+} to Mn^{2+} and the composion ratios of $H_3PO_4/HCLO_4$.

acid solution is put into a 100 ml. quartz beaker, next the calculated amount of $KMnO_4$ powder is added and then the whole system is heated. $KMnO_4$ will be decomposed at about 100°C and much O_2 will be expelled. If the solution temperature rises to 265° C, the heating source is taken away, the solution is cooled, diluted with water to 100 ml., and the extinction of the dilute solution is measured by using a 1 cm cell with 520 mμ wave length. The absorption curve and calibration curve are shown in Figs. 1 and 2 respectively.

1.2. Extent of Oxidation Method

Since sintered ferrites with definite quantities of Mn^{3+} and Fe^{2+} could not be made, samples having the mixed ratios as shown in Table 1 were used.

Table 1

ANALYTICAL SAMPLES (MOL %)

	S-①	S-②	S-③	S-④
Fe_2O_3	38	32	32	52
FeO		16	14	
$MnCO_3$	32	27		
MnO			29	30
ZnO	30	25	25	12
MgO				6

Sample S-4 is commercially available and the compositions of the sample are assumed as shown in Table 1 with the results of chemical analysis.

The dissolution of a sintered ferrite with an acidic aqueous solution means the destruction of a spinel structure, and divalent and trivalent ions of two different metals are produced at nearly the same time. The larger the difference of oxidation potential between the ions, the faster the reaction velocity of the electron donors and acceptors. Therefore, to determined quantitatively Fe^{2+} and Mn^{3+}

which are presumed to be in a sintered ferrite, it is necessary that the oxidation potential of the oxidizing solution should be higher than the oxidizing metalic ion(Mn^{3+}) in the solution, and that the dissolving time should be very short.

As a result of the above mentioned experiments, the author found an oxidizing solution satisfying the two necessary conditions: the mixed H_3PO_4-$HCLO_4$ solution series at high temperature. The composition of the dissolving solution is fixed at $H_3PO_4/HCLO_4=30/1$ from Fig. 3, and the experiments of decomposing temperature were studied from 255°C to 270°C. At such a high temperature, H_3PO_4 changes to poly phosphoric acid (the strong phosphoric acid) which has a very strong dissolved action.⁽¹⁾ Therefore, sample S-4 dissolved thoroughly in 10 seconds.

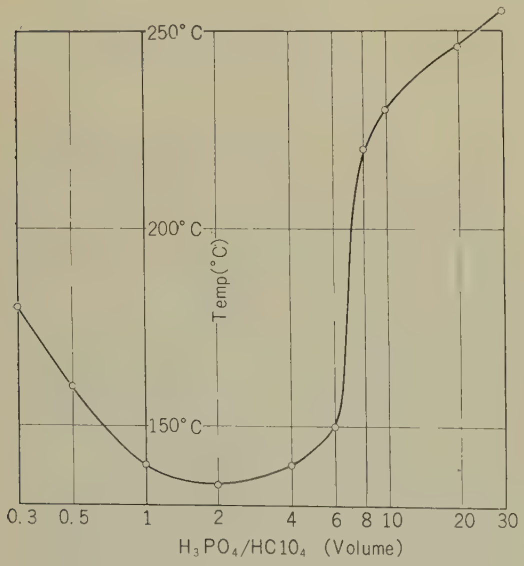


Fig. 3—Absorption curve of Mn^{3+} extinction.

This analytical method is named extent of oxidation method. The aim of the method is to study the Mn^{3+} in the sintered ferrites by statistical means. It is very important in this method to fix the analytical conditions. The conditions which must be fixed are as follows:

1. The heating conditions of the solution.
2. The temperature at which the sample is introduced into the solution.
3. The time interval after the sample is introduced until the heating source is taken away.
4. Whether the solution is to be stired or not after the heating source is taken, away.
5. The cooling location.
6. As the glass of mercury thermometer is corroded by heated poly phosphoric acid, it is necessary to minimize the errors due to the thermometer in the experiments.
7. It is necessary that the heating time of the solution into which the sample is introduced is kept almost constant throughout the experiments.

The orthogonal array of $L_{64}(4^{21})$ form was used to inspect these conditions.⁽²⁾ As a result, the best conditions were found to be as follows:

1. The heating conditions are: mixed acid is heated to about 175° C by gas for 5 minutes and then heated by a molten solder bath kept at 325±5° C up to the temperature at which the sample is introduced.
2. The most favorable temperature at which to introduce the sample is 265°C.
3. The most favorable time interval at which to take away the heating source after the sample is introduced into the solution is 10 seconds.
4. Neither stirring nor the cooling location (about 25°C) have any affect on the experimental results.

Moreover, the degree of dryness of the samples must be nearly constant for all the samples, and the quartz thickness of each beakers used must also be constant.

The statistical results of Table 1 are shown in Fig. 4.

1.3. Acceptability of Oxidation of Reduction Method

It is necessary to examine whether Fe^{2+} or Mn^{3+} is more rich in sintered ferrite. For this examination poly phosphoric acid containing Mn^{3+} is used.

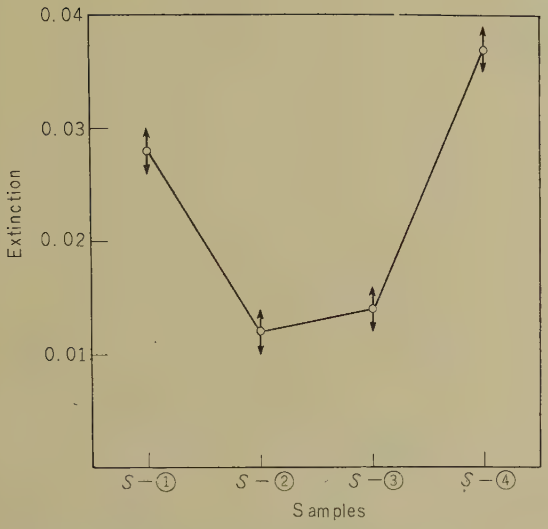


Fig. 4 ①—Main effect of samples. (Marks → show the values of the 95% tolerance limits, 0.002).

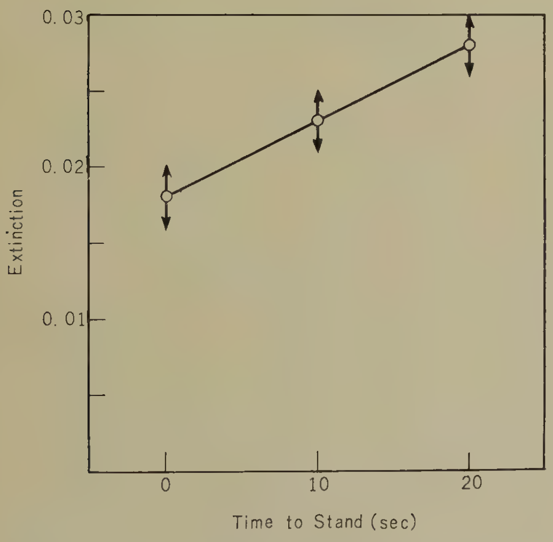


Fig. 4 ②—Main effect of reaction time. (Mark → show the values of the 95% tolerance limits, 0.002.)

The preparation of the solution as follows: The H_3PO_4 solution is heated, Mn_2O_3 powder is added to the heated solution at a temperature from 240°C to 250°C and decomposed; then the whole system is cooled and an amount of water equal to that which was evaporated during the above process is added. The poly phosphoric acid solution containing Mn^{3+} is the original solution.

The analytical results of Table 1 by this method are shown in Table 2. As a result, it becomes clear that the oxides containing Fe^{2+} and Mn^{3+} exist in the sample S-4 and that the amount of Fe^{2+} is more than that of Mn^{3+}.

Table 2

THE RESULTS OF "ACCEPTABILITY OF OXIDATION OR REDUCTION" METHOD (EXTINCTION)

S-①	0. 446	0. 450
S-②	0. 256	0. 248
S-③	0. 251	0. 246
S-④	0. 421	0. 420
Blank values	0. 428	0. 438

The analytical results of Table 1 by two different methods are as follows: FeO of sample S-2 is 7.5%, FeO, and Mn_2O_3 of sample S-3 are 8.6% and 0.1% respectively, and FeO and Mn_2O_3 of sample S-4 are 1.5% and 0.4% respectively. However, FeO and MnO contain small amounts of Fe_2O_3 and Mn_2O_3 respectively, because these oxides are reduced at high temperatures by H_2 gas. Also, the errors in estimating the amounts of FeO shown in Table 1 seem due to the different lots.

2. The Analysis of Iron Oxides, Manganese Oxides, and Two Different Mn-Zn Ferrites

2.1. Preparation of Samples

The author mixed Fe_2O_3, $MnCO_3$, and ZnO in guaranteed reagent in accordance with the composition ratios shown in Table 2, pressed the samples to a moderate size and sintered them two hours in air at 800° C.

Next, the sintered materials were powdered and pressed into pieces each weighing nearly 2 grams and these pieces were sintered under the conditions shown in Table 4. Each sintered sample of Table 3 was repeatedly examined three times under the same conditions,

Table 3

ANALYTICAL SAMPLES
(THE COMPOSITION RATIOS, MOL %)

	S-Fe	S-Mn	S-F-①	S-F-②
Fe_2O_3	100		50	55
$MnCO_3$		100	25	25
ZnO			25	20
			The composition ratios as to produce a spinel form.	The amount of Fe_2O_3 is 10 mol % exess over spinel form.

Table 4

THE SINTERING CONDITIONS

Sintered temp. (°C)	Quenched temp.			
	Treatments in N_2 (moderate cooling)	Treatments in air		
		1000° C	800° C	600° C
1360° C				
1200° C				
1000° C				

(The sintering time interval is 3 hours for each.)

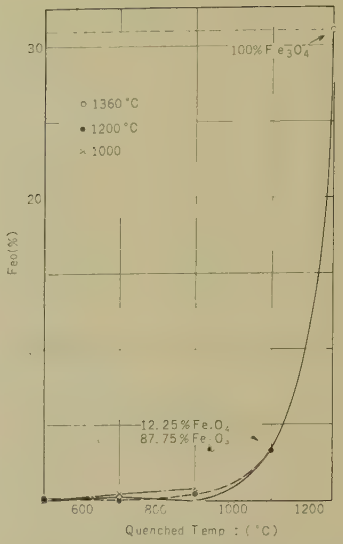


Fig. 5—Sintered iron oxide. The relations between the sintering temperature and FeO (%) for different quenching temperatures. (The value of 95% tolerance limits at measuring points are 0.4%).

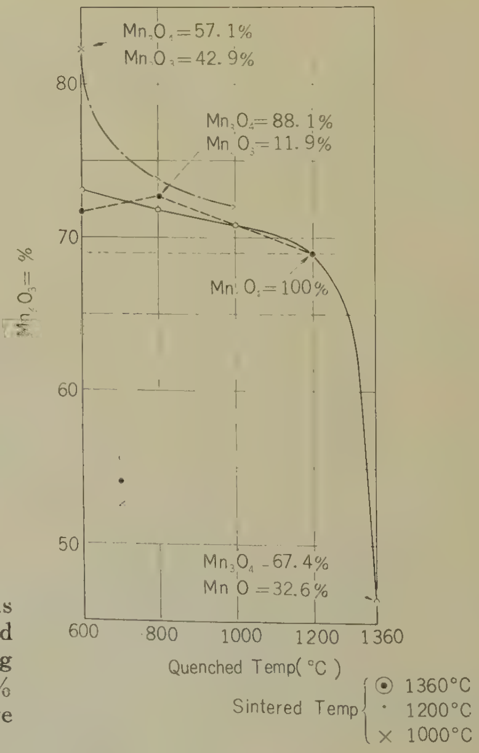


Fig. 6—Sintered maganese oxides. The relations between the sintering temperature and Mn_2O_3 (%) at different quenching temperatures. (The values of the 95% tolerance limits at measuring points are 1.5%).

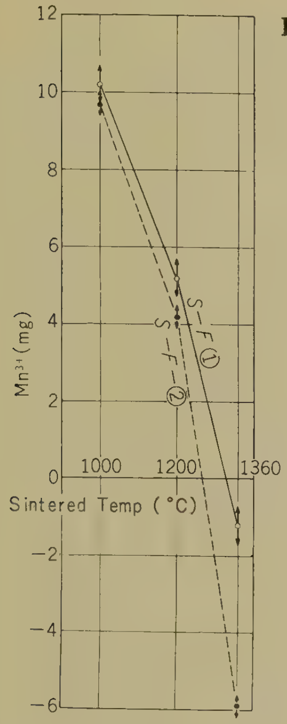


Fig. 7 ①—"Acceptability of oxidation or reduction" method. The relations between the sintering temperature and Mn^{3+} in the two ferrites. (Marks → show the values of the 95% tolerance limits.)

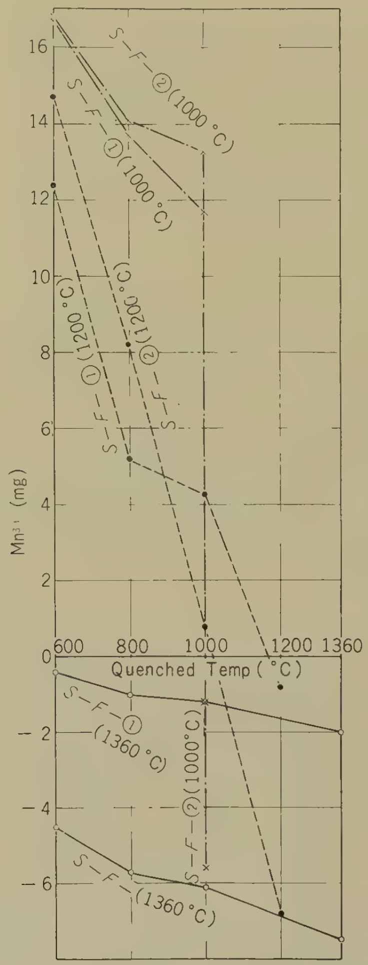


Fig. 7 ③—"Acceptability of oxidation or reduction" method. The relation between the quenching temperature and Mn^{3+} in the two ferrites which are sintered at 1000°C, 1200°C, and 1360°C. (The values of 95% tolerance limits at measuring points are 1.1 for S–F–1 and 0.6 for S–F–2 respectively.)

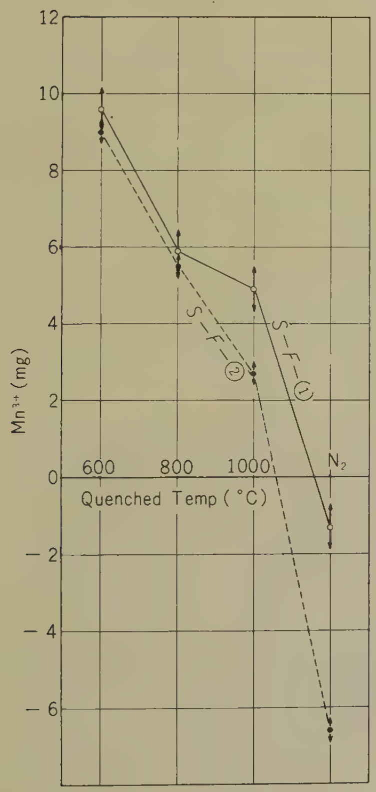


Fig. 7 ②—"Acceptability of oxidation of reduction" method. The relations between the quenching temperature and Mn^{3+} in the two ferrites. (Marks → show the values of the 95% tolerance limits.)

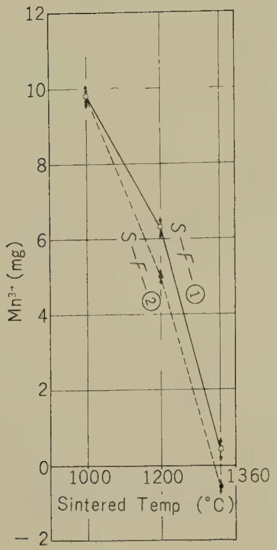


Fig. 8 ①—"Extent of oxidation" method. The relations between the sintering temperature and Mn^{3+} for the ferrites. (Marks → show the values of 95% tolerance limits.)

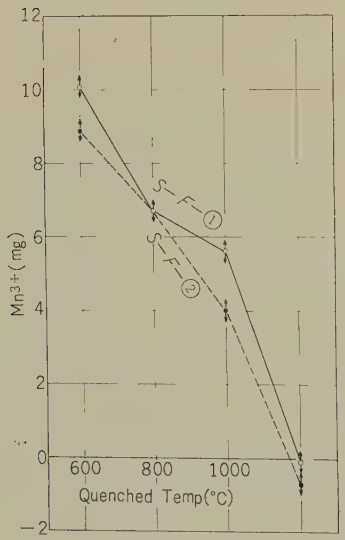


Fig. 8 ②—"Extent of oxidation" method. The relations between the quenching temperature and Mn^{3+} for the two ferrites. (Marks → show the values of 95% tolerance limits.)

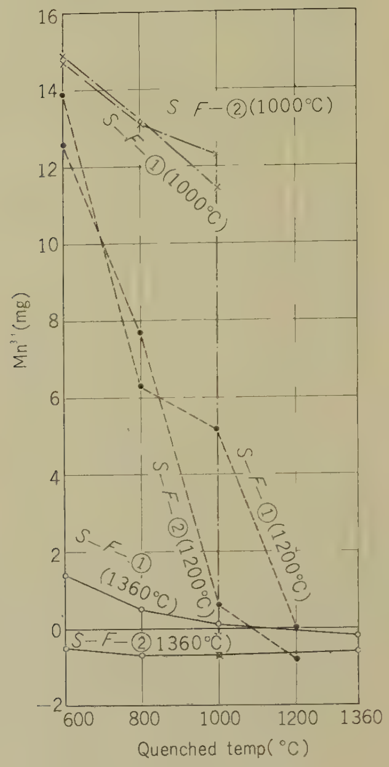


Fig. 8 ③—"Extent of exidation" method. The relations between the quenching temperature and Mn^{3+} for the two ferrites which are sintered at 1000°C, 1200°C, and 1360°C. (The values of the 95% tolerance limits at measuring points 0.6 for S–F–1 and 0.4 for S–F–2 respectively.)

and the results were statistically searched.

2.2. The Analysis of the Sintered Iron Oxidet

The results are shown in Fig. 5.

2.3. The Analysis of the Sintered Manganese Oxides

The results are shown in Fig. 6.

2.4.1. Acceptability of Oxidation of Reduction of Two Different Sintered Ferrites

The results are shown in Fig. 7 ①~③.

Table 5

ANALYTICAL RESULTS Fe^{2+}, Mn^{3+} IN THE TWO SINTERED Mn-Zn FERRITES

Sintered temp. (°C)	Quenched temp. (°C)	S-F-①		S-F-②	
		Fe^{2+} (mg)	Mn^{3+} (mg)	Fe^{2+} (mg)	Mn^{3+} (mg)
1360	Cooling in N_2	2.0±1.1	0 ±1.1	7.5±0.6	0 ±0.6
	1000	1.4	0.2	6.1	0
	800	1.6	0.6	5.7	0
	600	1.8	1.4	4.7	0
1200	Cooling in N_2	0.9	0.1	6.8	0
	1000	0.9	5.2	0	0.8
	800	1.1	6.3	0	8.2
	600	1.5	13.9	0	14.7
1000	Cooling in N_2	1.2	0	5.6	0
	1000	0	11.7	0	13.3
	800	0	13.7	0	14.1
	600	0	19.7	0	16.8

Table 6

OBSERVED RESULTS OF THE COMPOSITION IN S-F-①

Sintered temp. (°C)	Quenched temp. (°C)	$Zn_{1/2}Mn_{1/2}Fe_2O_3$ (%)	$Zn_{1/2}Fe_{1/2}Mn_2O_3$ (%)	FeO (%)	MnO (%)	Mn_2O_3 (%)	Fe_2O_3 (%)
1360	Cooling in N_2	97.4		1.3	13.		
	1000	98.9	0.3	0.8			
	800	98.3	0.9	0.8			
	600	97.3	2.0	0.7			
1200	Cooling in N_2	99.3	0.1	0.6			
	1000	95.0	2.4			2.6	
	800	94.0	2.8			3.2	
	600	88.1	3.8			8.1	
1000	Cooling in N_2	98.4		0.8	0.8		
	1000	83.1				8.4	8.5
	800	80.9				9.5	9.6
	600	75.8				12.0	12.2

Table 7

OBSERVED RESULTS OF THE COMPOSITION IN S-F-②

Sintered temp. (°C)	Quenched temp. (°C)	$Zn_{1/2}Mn_{1/1}Fe_2O_3$ (%)	Fe Fe_2O_4 (%)	Mn_2O_3 (%)	Fe_2O_3 (%)
1360	Cooling in N_2	87.5	10.2		2.3
	1000	87.5	8.3		4.3
	800	88.1	7.3		4.6
	600	87.2	6.1		6.6
1200	Cooling in N_2	87.5	9.2		3.3
	1000	86.0		0.5	13.5
	800	75.2		5.9	18.9
	600	65.8		10.9	23.6
1000	Cooling in N_2	87.5	7.6		5.0
	1000	68.0		7.6	22.5
	800	67.0		10.0	23.0
	600	62.8		12.1	25.1

2.4.2. Extent of Oxidation of Two Different Sintered Ferrites

The results are shown in Fig. 8 ①～③.

2.5. Analytical Results of Valencies in Two Different Sintered Ferrites and the Presumption of the Composition

The results are shown in Table 5, 6 and 7.

3. Conclusion

The method which is thought to be best in determining Mn^{3+} and Fe^{2+} in a sintered Mn-Zn ferrite was described. The results clearly showed that a sintered Mn-Zn ferrites contains small amount of oxides having Fe^{2+} and Mn^{3+}, and the analytical results of sintered materials of iron oxides and manganese oxides showed good agreement with those of other analytical methods.

Statistical analysis of the measurements on the synthetic samples and the sintered Mn-Zn ferrites showed that the velocity of Mn^{2+} oxidizing to Mn^{3+} is about one-tenth the velocity of Fe^{2+} oxidizing to Fe^{3+}. By considering the difference in oxidation velocity of these ions, the results shown in Table 5 were deduced. It is assumed that these results are reasonable.

Moreover, it is concluded that manganese oxide having tetravalent or higher valences does not exist from the observed experimental results.

Acknowledgement

The author wished to express his sincere appreciation to Y. Shichijo and T. Niimi for their guidance during this work.

References

(1) T. Kiba: "An analysis method using strong phosphoric acid." "*The latest analytical chemistry*" Vol. 7. Edited by the Kinki subdivision, *the Japan Socity for Analytical Chemistry*, p. 67, 1956.
(2) G. Taguchi: "Statistical experimental design," Vol. I and Vol. II, *Maruzen Co. Ltd.*, Tokyo Japan.

REVIEW OF THE ELECTRICAL COMMUNICATION LABORATORY
VOLUME 9, NUMBERS 5-6, MAY-JUNE, 1961

U.D.C 621. 317. 73 : [621. 385. 032. 2 : 621. 317. 335

An Input Admittance Meter for Electron Tubes*

Takuya KOJIMA† and **Hiroshi HARA**†

A study was made on a precision meter for measuring the input admittance of electron tubes at 70 Mc. Although the method of measurement is the common resonant-circuit method with a variable capacitor, this meter has the feature that the resonance point is obtained by detecting the phase angle of the resonant-circuit current. Compared with the ordinary method in which the point of maximum current amplitude is detected, the sensitivity of the present method is much higher and is very suitable for measuring the input capacitance of electron tubes, whose Q is low in the high frequency region. The schematic diagram of the meter, the method of measurement, measuring and experimental results are given: including experimental data on various electron tubes.

Introduction

As microwave communication equipment is increasingly used, demand has arisen for tubes not requiring readjustment of the intermediate frequency amplifier, which is a part of the equipment, when replacing tubes. Namely, formerly resonance was readjusted due to change in tube input and output capacitance every time electron tubes were replaced in the intermediate-frequency amplifier in equipment, but it is desired to eliminate the need for this adjustment in the future. For this purpose, however, the deviation in input and output capacitances of each electron tube should be kept within small limits. Furthermore, the center value of the capacitance distribution of many tubes acting on long communication networks should also be kept within an extremely small deviation. For example, as explained below it is necessary to maintain, the deviation of the center value of the input capacitance of the 6R-R8 to within 0.2 pF.

Tubes to meet this requirement can be produced by careful control during and after manufacture. Prior to such new production control, however, an instrument for measuring the input and output capacitance is first required. Therefore, an instrument for measuring input and output capacitances was developed.

The requirements for the measuring instrument are as follows:

1) It is necessary to measure input and output capacitances accurately:
Accuracy should be within 0.05 pF.

2) Frequency should be the actual frequency, 70 Mc.

3) Measurement should be conducted promptly and simply as the measurement of large quantities of tubes is required.

On studying measurement methods to meet these requirements, it was decided that a resonance method is most suitable from considerations of measurement frequency and accuracy. Input admittance of electron tubes contains a considerable conductance component at 70 Mc, and Q in the resonant circuit is low. For this reason, the method of measuring the phase of the circuit current,

---

* MS received by the Electrical Communication Laboratory, on November 2, 1960. Originally published in the *Kenkyu Zituyôka Hôkoku* (*Electrical Communication Laboratory Technical Journal*), Vol. 10, No. 5, pp. 969–979, 1961.

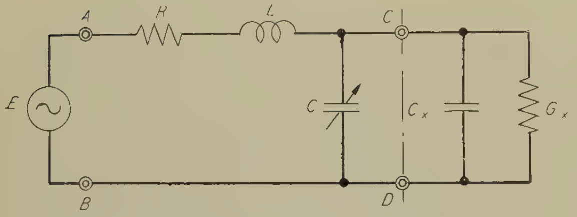
† Electronic Component Research Section.

which is a highly sensitive method of detecting the resonant point, has been adopted.

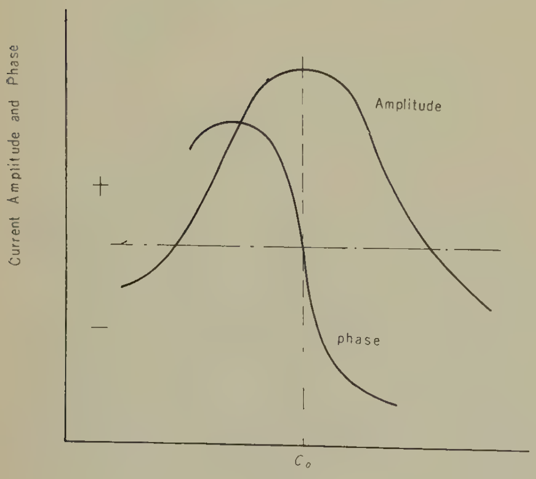
By the above process, accuracy in capacitance measurement was assured and measuring time could be shortened, thus resulting in the successful completion of a capacitance measuring instrument to meet the above requirements.

1. Description of Measuring Equipment

Fig. 1 shows the principle of the measuring equipment. The tube under test is connected in parallel with the variable capacitor C in the series resonant circuit at terminals C and D as shown in Fig. 1a. Resonance is adjusted by rotating the variable capacitor C when the tube under test is connected to the terminals and when it is not. The difference in readings of the variable capacitor gives the capacitance of the tube under test. There is nothing very different from the usual resonance method, except that in this case the point of resonance is detected by reading the phase of the current phase, but it is detected in the usual method by reading the point of maximum resonant-current amplitude. When measuring input capacitance, there exists comparatively large conductance in parallel with the input capacitance, thus lowering the value of Q of the resonant circuit. Therefore, as shown in Fig. 1b the current-amplitude resonance curve becomes less peaked, thus making it difficult to detect the point of peak current. However, phase varies quite rapidly at the resonant point; therefore it is considered easy to detect the point of resonance if phase is measured.



(a) Equivalent circuit



(b)

Fig. 1—Principle of resonant circuit.

An external view of the experimentally constructed admittance meter is shown in Fig. 2, and its schematic diagram is shown in Fig. 3. Two signals come out of the master oscillator at 70 Mc; one goes through the resonant circuit including the tube under test and the other through a phase shifter. The frequency of these two signals is converted to 455 Kc, but phase remains unchanged. These two signals are amplified by 455 Kc intermediate frequency amplifiers and applied to the vertical and horizontal axes of a cathode-ray oscilloscope. The master and local oscillators are crystal-controlled in order to stabilize the frequencies.

The operation of this equipment will be explained below for the case of input admittance measurements.

1.1. Measurement of Capacitance

First resonance is obtained without inserting the tube under test; Q of the resonant circuit is high and the resonant point as the point of maximum current amplitude is easily read on voltmeter V or determined by the point of maximum vertical-axis deflection on the cathode-ray oscilloscope shown in Fig. 3. In this condition the phase shifter of Fig. 3 is adjusted so as to make the Lissajou figure

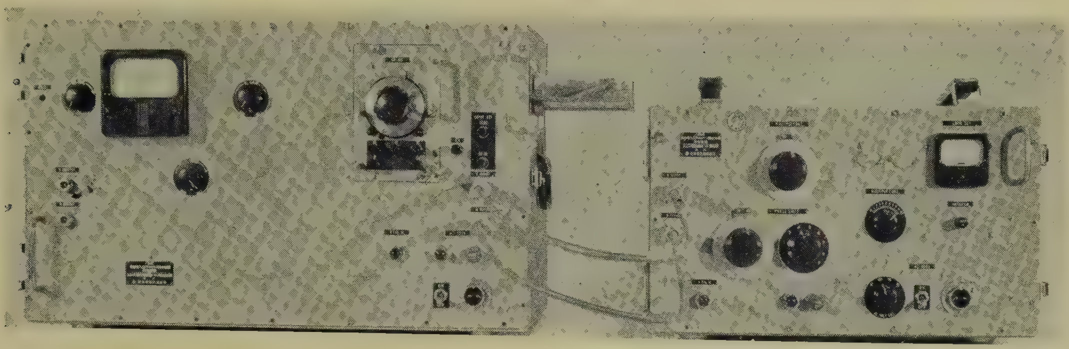


Fig. 2—Input admittance meter for electron tubes.

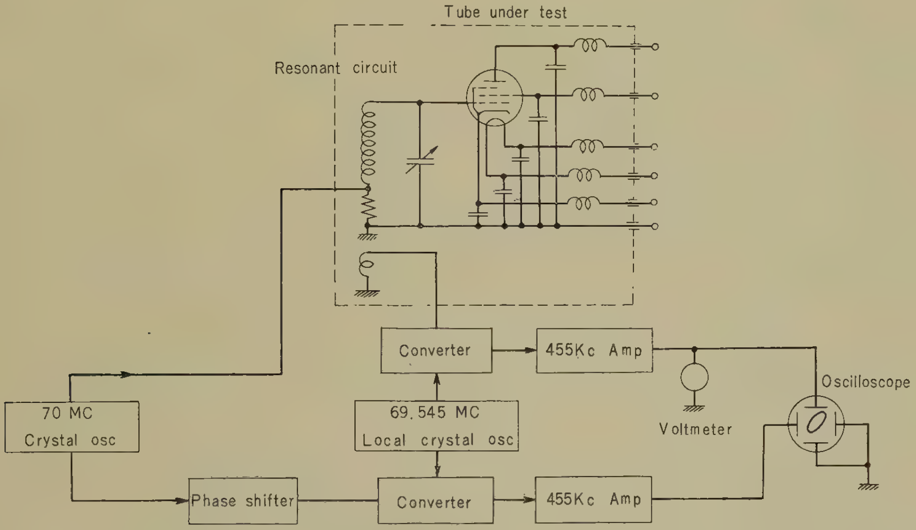


Fig. 3—Schematic diagram of input admittance meter.

a straight line.

Then the tube to be tested is inserted. To cancel the consequent increase in capacitance, the capacitance of the variable capacitor is decreased and the figure on the cathode-ray tube is again made a straight line. In other words, the variation in variable capacitor capacitance gives the capacitance of the tube under test.

1.2. Measurement of Conductance

The Q of the circuit can be measured by the half-resonance method. In the usual amplitude detecting method the half-resonance point is found by the fact that at this point the current amplitude becomes $1/\sqrt{2}$ times as large as that at resonance.

In our meter Q is obtained by observing the variation of the phase of the current with variation of the variable capacitor. The method will be explained in the order of the actual measurement.

Now let us assume that the capacitance measurement of the preceding paragraph has just been carried out; that is, resonance has been obtained with the tube-under-test inserted, and the figure on the cathode-ray tube remains a straight line. From this condition, the measurement of the conductance component will be carried out.

In the first place the phase shifter dial is rotated and the phase of the comparison circuit is varied by a fixed angle 1. The figure on the cathode-ray tube, of course, becomes elliptic.

Then, by rotating the variable condenser of the resonant circuit, the figure on the cathode-ray tube is adjusted so as to again become a straight line. Conductance can be obtained from a chart which gives the relations between the value of capacitance variation and that of conductance. A fixed phase angle such as 45°, 30°, etc., is used.

2. Principle of Measurement

The operating principle of the measuring instrument has been shown in Fig. 1. As stated in paragraph 1, the feature of this measuring instrument lies in the detection of the phase angle of the resonant current. Admittance Y between terminals A and B of Fig. 1 is given as follows.

$$Y=\frac{RG_x^2+G_x+\omega^2R(C+C_x)^2+j\omega\{(C+C_x)-LG_x^2-\omega^2L(C+C_x)^2\}}{\{RG_x+1-\omega^2L(C+C_x)\}^2+\omega^2\{R(C+C_x)+LG_x\}^2} \qquad (1)$$

where

L: inductance of coil of resonant circuit
R: resistance of coil
C: capacitance of variable capacitor
C_x: capacitance of tube under test
G_x: conductance of tube under test.

Therefore the phase angle of the current with respect to the voltage is

$$\tan\phi=\frac{\omega\left\{(C+C_x)-\frac{G_x^2}{\omega^2C_0}-\frac{1}{C_0}(C+C_x)^2\right\}}{RG_x^2+G_x+\omega^2R(C+C_x)^2} \qquad (2)$$

where $C_0=1/\omega^2L$, and C_0 is the capacitance of variable condenser at resonance when no tube is inserted.

2.1. Measurement of Capacitance

The procedure of capacitance measurement in paragraph 1 is reviewed here. The point of maximum current amplitude when there is no tube, i. e., G_x 0, may readily be obtained from

(1) and occurs for the following capacitance:

$$C+C_x=C_0 \qquad (3)$$

and the current phase ϕ is

$$\phi=0 \qquad (4)$$

In this case the point of maximum current amplitude coincides with the point of zero phase difference.

Next the variable capacitance is adjusted so as to make current phase zero again after inserting a tube. The point of zero current phase difference in this case occurs for the following value of the capacitance:

$$C+C_x=C_0\left(\frac{1}{2}+\sqrt{\frac{1}{4}-\frac{G_x^2}{\omega^2C_0}}\right)$$

$$\fallingdotseq C_0\left(1-\frac{G_x^2}{\omega^2C_0}\right) \text{ where, } G_x\ll\omega C_0 \qquad (5)$$

As shown in Eq. 5, the measured value of $C+C_x$ is affected by G_x. In the practical range, however, the error in $C+C_x$ caused by G_x is small and negligible, as will later be clarified by calculation of actual values. Furthermore, the value of $C+C_x$ of the maximum current amplitude is given.

$$C+C_x=C_0\left\{\frac{1}{2}-RG_x-\sqrt{\left(\frac{1}{2}+RG_x^2\right)+\frac{G_x^2}{\omega^2C_0^2}}\right\} \qquad (6)$$

2.2. Accuracy in Capacitance Measurement

Sensitivity of detection of current phase at the point of zero phase angle determines the accuracy of capacitance measurement of this measuring instrument.

Measurement sensitivity S at the point of zero current phase will be represented by the following formula. Using ϕ which can be evaluated from Eq. 2,

$$S=\left\{\frac{d\phi}{d(C+C_x)}\right\}_{\phi=0} \qquad (7)$$

and when Eq. 2 is substituted, S will be obtained.

When ϕ is small,

$$\tan\phi=\phi$$

can be used and

$$S=\omega\frac{1-2\dfrac{C+C_x}{C_0}}{RG_x{}^2+G_x+\omega^2R(C+C_x)^2}\,. \qquad (8)$$

The results of the calculation of Eq. 8 will be given later in the paragraph on experimental results. If, however, an ellipse of 1° is distinguishable from a straight line, it is understood from the above equation that a measurement accuracy of 0.02 pF is obtainable when conductance G_x is small.

2.3. Measurement of Conductance

After resonance is obtained with the tube under test connected and the figure on the cathode-ray tube becomes a straight line, the phase shifter is operated and the phase of the comparison current is varied by a fixed angle $\Delta\phi$. Then the figure on the cathode-ray tube is again made a straight line by rotating the variable capacitor. These procedures are shown by the following expression:

$$\tan\varDelta\phi=K \qquad (9)$$

then, from Eq. 2:

$$K=\frac{\omega\left\{(C+C_x)-\dfrac{G_x{}^2}{\omega^2C_0}+\dfrac{1}{C_0}(C+C_x)^2\right\}}{(RG_x+1)G_x+\omega^2(C+C_x)^2R} \qquad (10)$$

To solve Eq. 10 for $(C+C_x)$

$$C+C_x=C_0\left[\frac{1}{2(1+K\omega C_0R)}+\sqrt{\frac{1}{4(1+K\omega C_0R)^2}-\left\{\frac{G_x{}^2}{\omega^2C_0{}^2}+\frac{KG_x}{\omega C_0(1+K\omega C_0R)}\right\}}\right] \qquad (11)$$

When the figure on the cathode-ray tube is a straight line before adjusting the phase shifter, it means $K=0$, and the value of $C+C_x$ is

$$(C+C_x)_0=C_0\left\{\frac{1}{2}+\sqrt{\frac{1}{4}-\frac{G_x{}^2}{\omega^2C_0{}^2}}\right\}. \qquad (12)$$

Then, the capacitance value, when the figure on the cathode-ray tube is made a straight line again after adjusting the phase shifter, is

$(C+C_x)_K=$ the right-hand side of Eq. 11. (13)

The difference $\varDelta C$ between Eqs. 10 and 11 is a function of K and G_x, and is shown by:

$$\varDelta C=(C+C_x)_K-(C+C_x)_0$$

$$=C_0\left[\frac{1}{2(1+K\omega C_0R)}+\sqrt{\frac{1}{4(1+K\omega C_0R)^2}-\left\{\frac{G_x{}^2}{\omega^2C_0{}^2}+\frac{KG_x}{\omega C_0(1+K\omega C_0R)}\right\}}-\frac{1}{2}-\sqrt{\frac{1}{4}-\frac{G_x{}^2}{\omega^2C_0{}^2}}\right] \qquad (14)$$

Thus it may be well to keep K constant, to make a chart of the relation between $\varDelta C$ and G_x, and to obtain G_x from $\varDelta C$.

3. Consideration on Experimental Results

In paragraph 2 the principle of measurement of this measuring instrument was stated. In view of the fact, however, that the frequency of measurement is high as 70 Mc and that high measurement accuracy is required, it is found necessary to confirm by experiment that the principle stated in paragraph 2 will be satisfactory in practice. Thus, as shown by the construction sketch in Fig. 4, tube-under-test models were prepared which were constructed with 9-pin miniature tube bases on which 10 pF ceramic capacitors and carbon resistors of various sizes (2000, 1000, 700, 600 and 500) were connected in parallel. The values of capacitance and conductance are calibrated in the order to be described below. It is possible, however, that the capacitance and conductance values of such test models at 70 Mc may be considerably changed by load inductance effects from those measured at low frequency. To check this point, a simple calculation was worked out, as shown in the Appendix. It was found by the result of this calculation that the effect of load inductance is negligibly small.

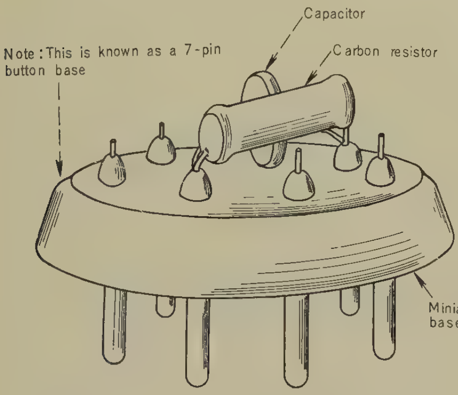


Fig. 4—Capacitor and resistor mounted on tube base.

3.1. Consideration on Capacitance Measurement

As shown in Eqs. 5 and 6 the measured value of capacitance is effected by the conductance G_x, whether the amplitude method or the phase method is used. Fig. 5 shows how the error occurs in actual practice. Curve 1 in the figure shows Eq. 5 and curve 2 Eq. 6. Parameters ( ) are as follows:

$$\left.\begin{aligned} C_0&=43.41\ \text{pF} \\ R&=1.5\Omega \end{aligned}\right\} \qquad (15)$$

As found in Fig. 5, when G_x is large, deviation of the measured capacitance value from its true value will increase for both methods. These deviations were experimentally examined with the test models mentioned above. As, however, it is difficuit to determine the true value of their capacitance, the same models were measured by the amplitude method and by the phase method. Then the difference between the two measured values and that between curves 1 and 2 in Fig. 5 were compared.

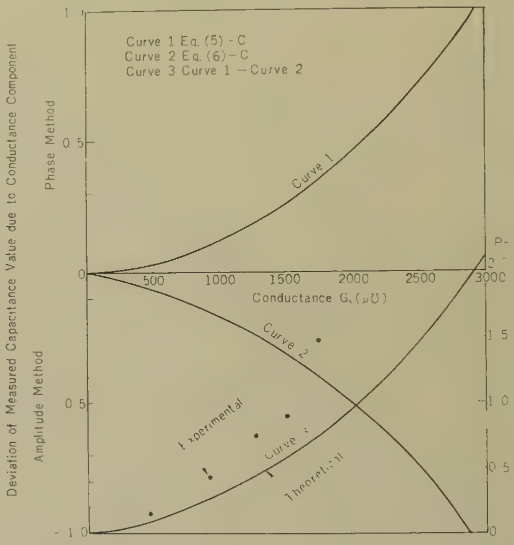


Fig. 5—Deviation of measured capacitance value due to conductance component.

Curve 3 in Fig. 5 shows the difference between curves 1 and 2 and points on the figure show the difference between values measured both by amplitude and phase methods. It is thus clarified, though indirectly, that good agreement between the theory and experiment is obtained.

Next the sensitivity of the capacitance measurement will be considered. The sensitivity of this measuring instrument is determined by the minimum distinguishable phase difference of the figure on the cathode-ray tube and also by the equivalent capacity variation. However, if G_x is large, variation

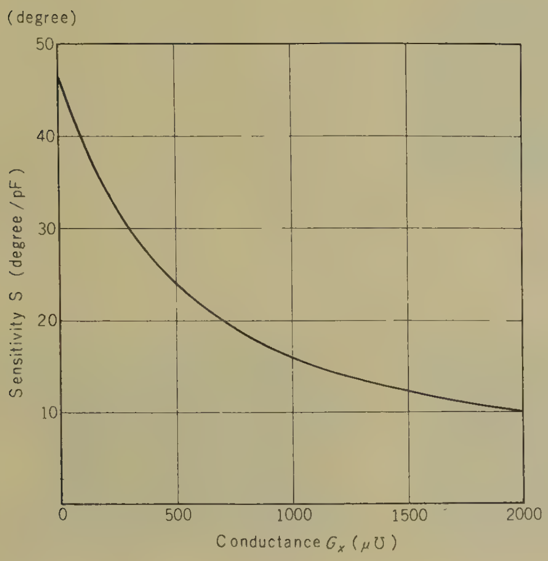
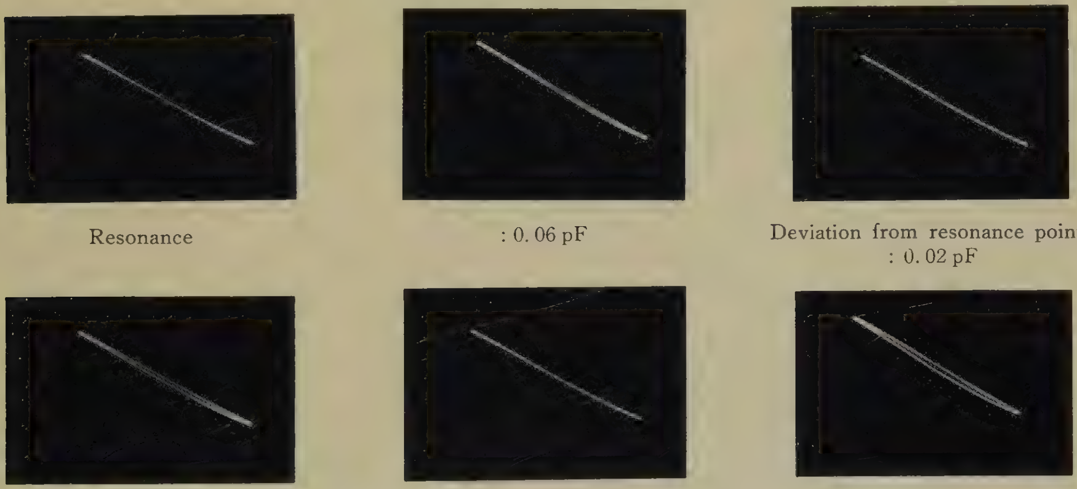


Fig. 6—Sensitivity of input admittance meter.

in current phase becomes small and sensitivity decreases. This calculation has already been carried out in Eq. 8. Using the actual value of the components of the instrument which are given in Eq. 15, Eq. 8 is plotted as Fig. 6. As seen in the figure, sensitivity S is 46°/pF when $G_x=0$. It has been ascertained by experiment that 0.02 pF was distinguishable when conductance was small. In Fig. 7 a photograph of an example of the pattern on the cathode-ray tube during the sensitivity measurement is shown.

3.2. Considerations on the Conductance Measurement

Conductance measurement was carried out in the order described in paragraphs 2 and 3. Eq. 14 gives the relations between the fixed angle phase shift of the phase shifter



Resonance

: 0. 06 pF

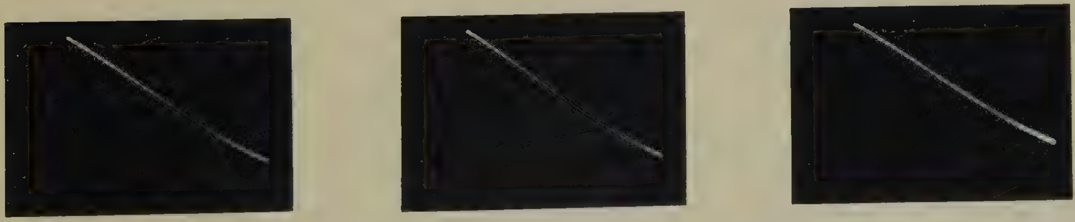
Deviation from resonance point
: 0. 02 pF

: 0. 08 pF

: 0. 04 pF

: 0. 1 pF

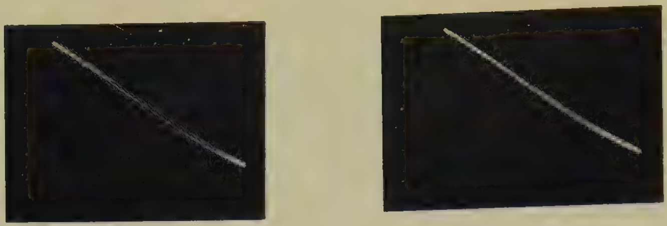
(a) for $G_x=0$



Resonance

: 0. 08 pF

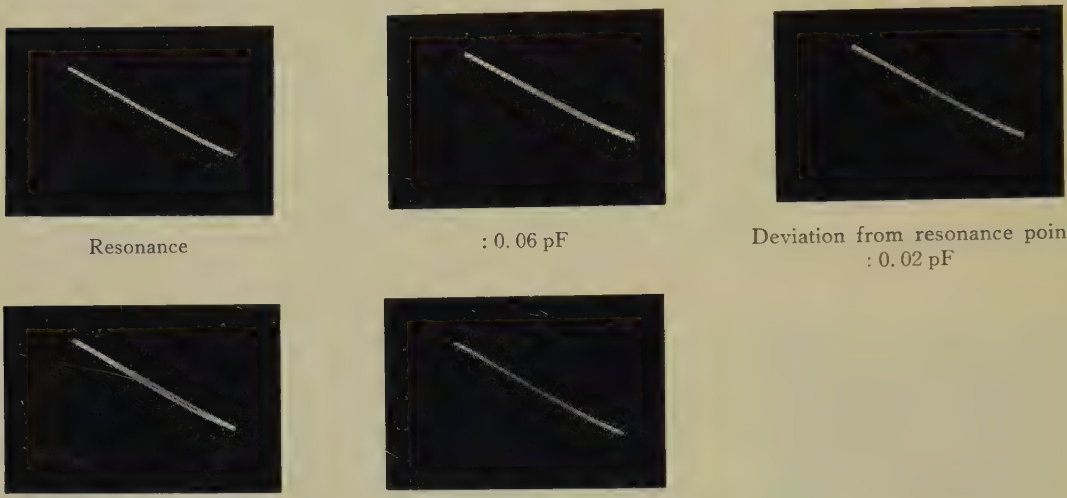
Deviation from resonance point
: 0. 04 pF



: 0. 1 pF

: 0. 06 pF

(b) for G_x=500 μ℧



Resonance

: 0. 06 pF

Deviation from resonance point
: 0. 02 pF

: 0. 08 pF

: 0. 04 pF

(c) for G_x=1000 μ℧

Fig. 7—Measured Sensitivity.

and the variation of the capacitance of the variable condenser necessary to make the pattern on the cathode-ray tube figure a straight line again. Eq. 14 is shown in Fig. 8, with the constants of Eq. 15 and $K=-1$ ($\Delta\phi=45°$) substituted. Now, as described in the beginning of this paragraph, various models of the tube under test are inserted and their measured values of conductance are plotted in the figure. The experimental results agree with the calculated curve.

3.3. Measurement of the Input Admittance of Electron Tubes

Last some examples of measurement of input admittance of electron tubes will be

Table 1

MEASURED VALUES OF INPUT CONDUCTANCE OF VARIOUS TUBES

Tubes	Conductance (μ℧)	
	Maker A	Maker B
6R-R8 (broad band pentode)	440~480	340~380
6R-R8C (broad band pentode)	440	
6R-R21 (broad band tetrode)	610~700	

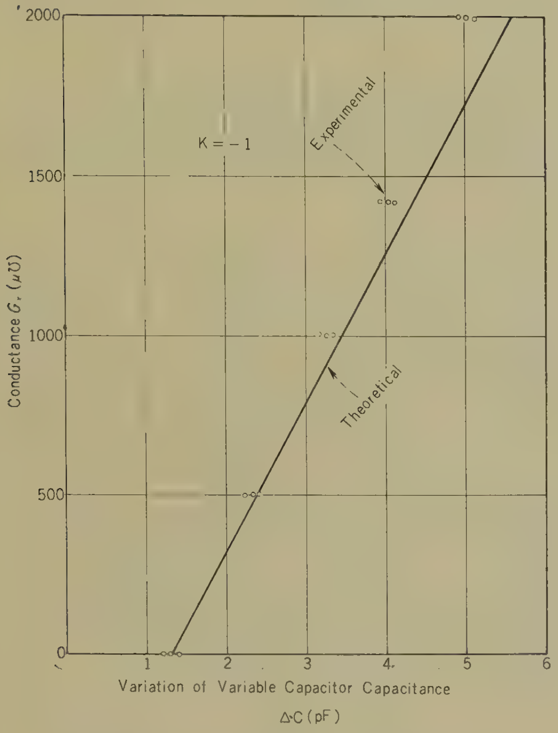


Fig. 8—Conductance measurement.

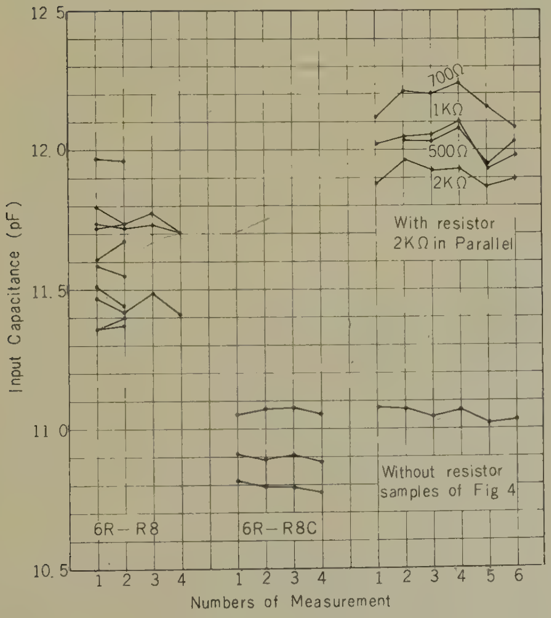


Fig. 9—Measured results for electron tubes and samples of Fig. 4.

given. This experiment was conducted to confirm the reproducibility of the measured values. The results are shown in Fig. 9 and show that the reproducibility of the measured values is satisfactory. The measurements were conducted at intervals of several days to several months. During this period no readjustments of the measuring instrument were made. In actual practice, however, several kinds of standard capacitors will be prepared and the reading of variable condenser of the measuring instrument will be corrected with these standard capacitors before measuring tubes. The measured values of input conductance at 70 Mc of various tubes are given in Table 1.

Conclusion

An input admittance meter for electron tubes has been developed. The principle of measurement consists of the common resonance circuit method with a variable capacitor. Before and after the insertion of the tube under test the circuit is resonated by rotating the variable capacitor. The tube capacitance can be determined from the difference of the readings of the variable capacitor. It is a feature of this meter that the resonance point is detected by measuring the phase of the current, and this method is very effective when the Q of the resonant circuit decreases due to the existence of a relatively large conductance component. In this way, measuring equipment having measurement accuracy to 0.02 pF has been realized.

A method of measuring the conductance component by observation of the phase of the resonance current was also given.

The construction of the equipment is outlined and the principle of measurement as well as the results of experiment are described together with some examples of input admittance measurements on various electron tubes.

The author has discussed only the problems of input admittance measurement. The output admittance, however, is measured in exactly the same way.

The instrument was manufactured experimentally by the Ando Denki K.K., and the cooperation of their staff is much appreciated.

Thanks are also due to Dr. M. Kuwata, chief of the Electronic Parts Research Section.

Appendix

The actual admittance of tube-under-test models, in which a carbon resistor and a ceramic capacitor are mounted on tube bases as shown in Fig. 4, it is feared, will be different because of lead inductance from the indicated values of the resistor and capacitor. Particularly in this experiment special attention should be paid to the value of the resistor, as mentioned in the paper. Now by presuming the lead inductance of the test tube model will be calculated.

The equivalent circuit of the tube-under-test model is considered to be as shown in Fig. 1, where L_1 shows the inductance of the lead of the base and L_2 and L_3 the lead inductances of the resistor and capacitor, respectively. The admittance Y between the terminals of this circuit is represented by the following formula:

$$Y=\frac{\{1-\omega^2C(L_2+L_3)+j\omega RC\}\,[R\{1-\omega^2C(L_1+L_3)\}-j\omega\{L_1+L_2-\omega^2C(L_1L_2+L_2L_3+L_1L_3)\}]}{R^2\{1-\omega^2C(L_1+L_3)\}^2+\omega^2\{L_1+L_2-\omega^2C(L_1L_2+L_2L_3+L_1L_3)\}^2} \qquad \text{(Ap. Eq. 1)}$$

Now, the values of the lead inductance of the stem, resistor and capacitor, and also of capacitance C, estimated from the structure, are assumed to be as follows:

$$L_1=1.5 \quad 10^{-8}\ \text{H}$$

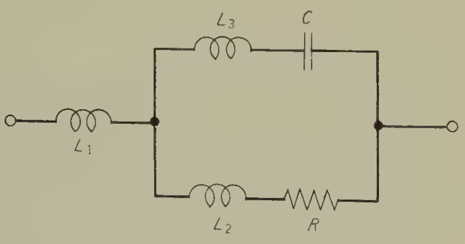
$$L_2=L_3=7. \quad 10^{-9}\ \text{H}$$

$$C=10\ \text{pF}$$

and if Y is worked out by substituting these values to Ap. Eq. 1:

$$Y=\frac{0.932R+)(0.00421R^2-0.921)}{0.918R^2+89.1} \qquad \text{(Ap. Eq. 2)}$$

From this formula it was found that conductance component of Y, when $R=500\sim2000\,\Omega$, gave an approximate error of less than 1% compared with $1/R$.



Ap. Fig. 1—Equivalent circuit of samples.

* * * *

U.D.C. 621. 391. 812. 6 : 538. 566
621. 371. 001. 2

Characteristics of Fading due to Ducts and Quantitative Estimation of Fading*

Sigetaka UGAI†

As a result of long-term propagation tests carried out on four frequencies simultaneously, it has been clarified that microwave fading due to the irregular distribution of the refractive index of the air has two modes of variation—long-period variation and short-period variation. The characteristics of each have been described.

The quantitative relation between these two variations was found experimentally. It has been brought to light after considerations by geometrical optics and experimental analysis that the long-period variation is closely related to the statistics of the distribution of the refractive index of the air, which can be obtained by radiosonde observation. On this basic applications of radiosonde data for the method of estimating the fading ranges in UHF and SHF bands for all seasons and various places is described.

Introduction

Microwave multiple-relay circuits which make up a network must have very high stability and very high transmission reliability. Fading phenomenon is inevitable in wireless transmission, and the cause for it is related to such complicated distributions of the refractive index of the atmosphere that very few researches have been made for the purpose of design for practical radio communication.[1) 2)] For the 4,000 Mc band, the characteristics of fading have been clarified to some extent by classifying them into k-type fading and duct-type fading. The fading mechanism of one of these types of fading, k-type fading, can be physically understood; therefore, to some degree quantitative estimates of the fading on other bands is possible.

As for duct-type fading, its production is due to such complicated vertical distribution of the refractive index of the air that the atmospheric conditions cannot be observed with ease, and this has deterred exact grasp of the fading characteristics. In Japan, because of the peculiar topography, a grasp of the general fading characteristics is the more necessary for designing circuits.

Formerly, in case of the quantitative estimation of duct-type fading in the 4000 Mc band, the value of mean vapor pressure on the ground was used[3)] as a radio-meteorological factor.

The value of the vapor pressure on the ground, however, is not sufficient to physically explain the fading mechanism. Practically, fading does not always correspond to it with the change of season. Moreover, it has been found that the value of fading obtained by the estimation formula[3)] tends to become much larger than the measured value when the distance is 100 km or longer. Therefore, the authors have made researches on the relation between fading characteristics and

---

* MS in Japanese received by the Electrical Communication Laboratory on 26 March, 1959. Originally published in the *Kenkyû Zituyôka Hôkoku* (*Electrical Communication Laboratory Technical Journal*), N.T.T., Vol.8, No. 7, July pp. 1003–1038, 1960.

† Radio Propagation Research Section.

radio-meteorology giving special effort to the radio-meteorological observation.

As a result of long-term experiments with more than two frequencies made simultaneously over the same propagation path, it has been found that the fading sequence has the two types of variation, long- and short-period variation.

Besides, it has been found statistically through the characteristics that the long-period variation is closely related to the variation of the mean gradient of the refractive index. In this way, the dependence of fading on frequency, on distance, on season, on region, etc. have experimentally ascertained, and besides a method of estimating fading range vs frequency over wider band (UHF- and SHF-band) has been found. Therefore in this paper, the description of the general characteristics of fading will be made on the basis of the experimental results.

1. Observed Characteristics of Fading

1.1. The Definition of Fading Treated in This Research

When a radio relay station for a microwave relay circuit is set up, it is important to take into account the topography along the propagation path in order to reduce the fading due to reflected waves from the ground as much as possible.

However, what causes the greatest amount of disturbance to communication is thought to be the deep duct type fading.

So-called duct-type fading was considered due to radio duct, i. e. negative gradient in the modified refractive index (M) of air. But it can be explained approximately by geometrical optics[2)3)4)], that the fade-out appears due to not only the negative gradient but also due to the existence of a bending part in the vertical distribution of the modified index, even when there is no radio duct. Thus fading can appear, even if there is no radio duct, when the shape of the M profile varies with time. Concerning this, description will be made later.

At still higher frequencies, not only the loss due to the above-mentioned fading but also the loss due to the absorption and scattering effects of oxygen, vapor, fog, and rain in the atmosphere can never be ignored. Among them, the loss by the effect of rain is largest. Here, however, these effects are omitted. This paper deals with only field intensity variation due to the shape of the irregular distribution of the refractive index in the air. Research has been made in persuit of its general characteristics with this restriction.

In other words, the variation of field intensity due to the influence of the ground (reflected wave, scattering, and diffraction on the ground) is not dealt with.

Later in this paper, "fading" means the field variation in the above-stated range.

1.2. Outline of Experiment on Propagation

1.2.1. Propagation Path

In research on the general characteristics of fading (the effects of propagation distance, path inclination to the ground, mean path height, season, etc.), it is not so easy as in case of research on other physical and chemical fields to classify the factors by making them uniform artificially, because nature has many factors that cannot be controlled.

Roughly speaking, the factors affecting fading characteristics are as follows:

1) Effect of distance
2) Effect of season
3) Effects of the refractive index gradient along the propagation path and the mean path height above the ground.

In the experiments it is most desirable to make some effects of these factors as equal as possible to each other, and to grasp the characteristics affecting the fading as exactly as possible under the given conditions.

Usually, the atmosphere is considered to have a stratified structure nearly parallel to the surface of the ground. As a result, fading is expected to appear most remarkably when a propagation path runs parallel with the ground surface (Section 2). In this paper,

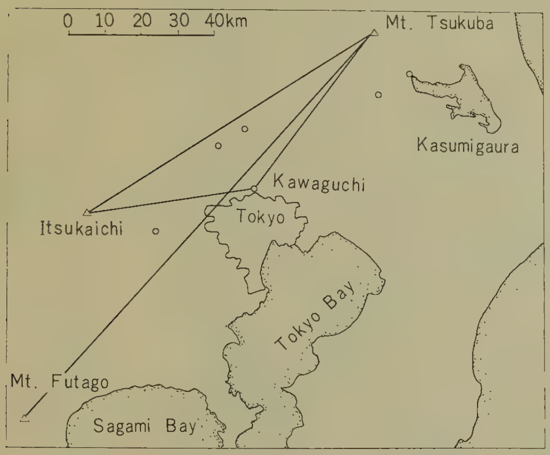


Fig. 1—Map of propagation paths.

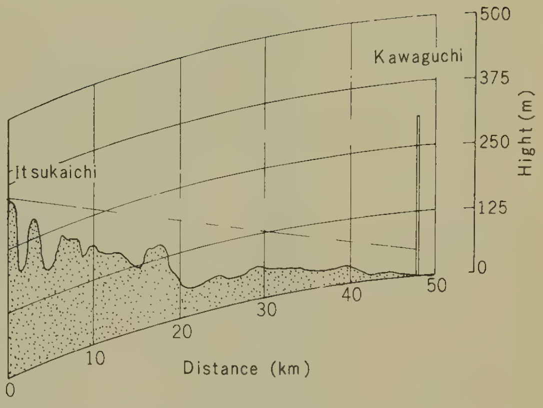


Fig. 2 (**c**)—Path profile from Kawaguchi to Itsukaichi. 4/3 earth's radius.

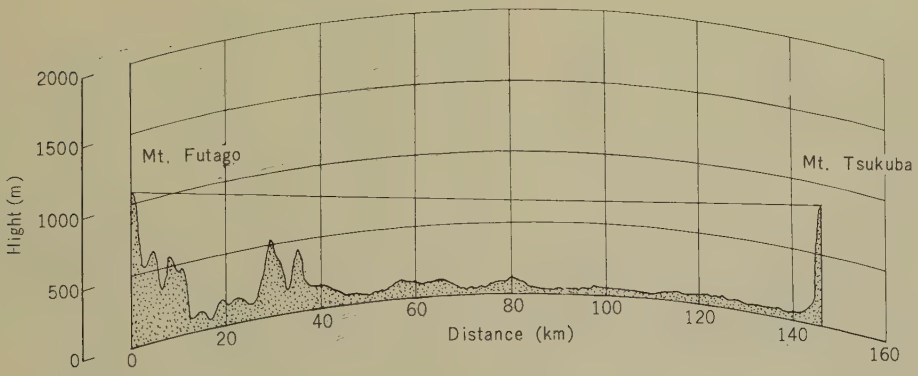


Fig. 2 (**a**)—Path profile from Mt. Futago to Mt. Tsukuba. 4/3 earth's radius.

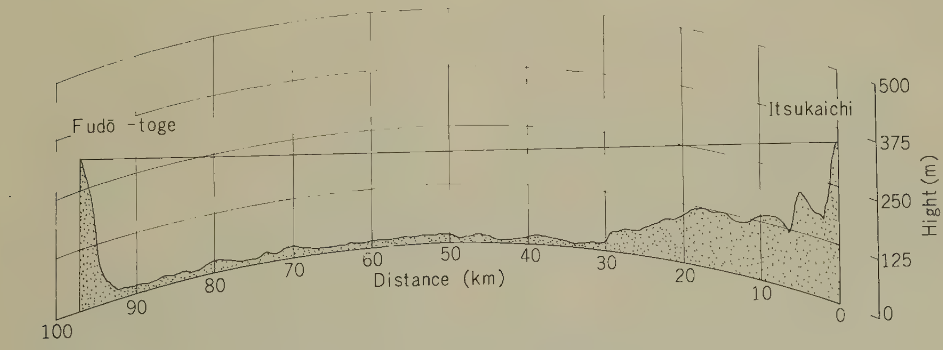


Fig. 2 (**b**)—Path profile from Fudô-tôge to Itsukaichi. 4/3 earth's radius.

Table 1

CONSTANTS IN PROPAGATION PATH

Propagation path	Height (m)		Path length (km)	Frequency (Gc)
	Trans.	Rec.		
Mt. Futago -Mt. Tsukuba	1090	870	147	0. 2 1. 5 4. 02 6. 0
Fudo-tôge -Itsukaichi	320	340	97	0. 2 1. 5 4. 02 6. 72
Kawaguchi -Itsukaichi	53	340	48	1. 5 4. 02 6. 72

only the fading due to nonlinear refractive index distribution is dealt with. Therefore the establishment of the experimental circuits was made under the following conditions:

(1) Path is parallel to the ground surface.

(2) Reflected wave on the ground is shielded by topographical features or objects on the ground.

(3) Long-term correspondence of fading to the refractive index vertical structure can be obtained.

To investigate the dependence on distance, the experimental circuits in Fig. 1 were selected.

Fig. 2 and Table 1 show respectively profiles of the experimental circuit and their constants.

1.2.2. System for Measuring and Recording Fading

The received power due to propagation is shown by the following relation:

$$P_r = P_t \cdot G_t \cdot G_r / L_t \cdot L_r \cdot A, \qquad (1)$$

where P_t, (P_r) = Transmitted (Received) power
G_t, (G_r) = Transmitting (Receiving) antenna gain
L_t, (L_r) = Transmitting (Receiving) feeder loss
A = Propagation loss

If the values of $P_t, G_t, G_r, L_t,$ and L_r are constant, the variation of P_r is determined by that of A. The value of A is determined by the topography of the propagation path and the spatial distribution of refractive index of the air along the path. The value of A cannot be measured directly, and so it was calculated through Equation (1) by measuring P_r and replacing it with the output from a standard signal generator whose output was known.

To get the records of rapid fadings, suitable recording systems were used. A 5-mA recording ammeter was used for the fading speeds less than 1 c/s, a pen writing oscillograph for those less than 20 c/s, and an electromagnetic oscillograph for these less than 1000 c/s were used.

The long-term measurement, e. g. the measurement of seasonal variation, was mechanized, and a special automatic level meter was used.

Fig. 3 shows the equipment. This consists of a level-reading device, a computer, an electric clock, and a camera to take photographs of the level count at a certain regular time. The sampling speed was 1/15 $\sec^{-1}$.

1.3. Statistical Expression of Fading

There are various methods to show quantitatively the value of received power, which varies with the time. One method that is in common use is the method by showing the ratio of the maximum power to the minimum power, in dB, at every given sampling time interval.

By this method, the treatment is comparatively easy; therefore, this is very suitable for the relative expression of the long-term variation. In this paper, it is expressed by the notation dB (max., min./5 min.), for example, instead of dB (maximum, minimum per 5 minutes). This is used for the expression of the hourly and diurnal characteristics of the received power.

Another method, the most ideal one, is as

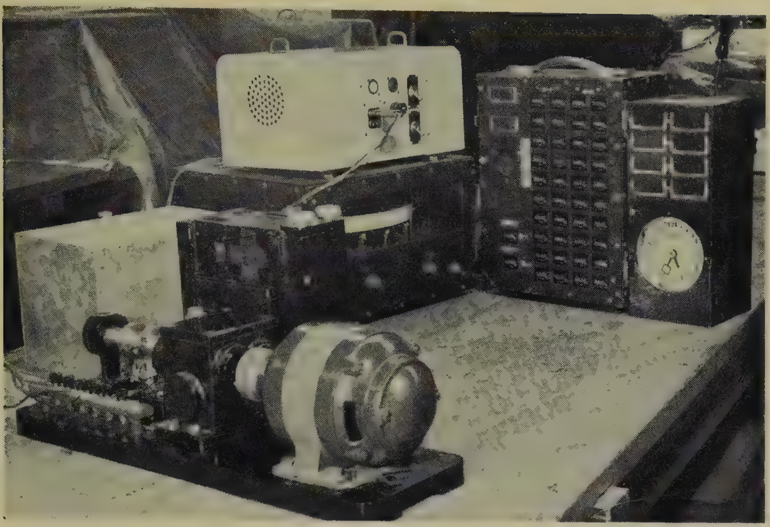


Fig. 3—View of automatic level counting device.

follows: the range of power variation is divided into discrete power levels, and then the total of the time length when the received power falls below the respective level is obtained. Next, the percentage of the total time length to the whole measuring time length is shown for every level. This indication corresponds to the limit of an infinitesimal sampling interval.

In this case, $f(z)$, the probability density function of z is expressed by

$$\text{Prob}\,(Z\leq Z_p)=\int_0^{z_p} f(z)dz=\int_0^{p} dF(z)=P$$

Accordingly, the value of received power, Z_p, can be determined by designating probability P. Especially, Z_p for $P=0.5$ is the median value of the received power in the period. The median value nearly corresponds to the free space value calculated for the line-of-sight propagation path. The deviation between the median value and calculated free space value is within 2 to 3 dB. Fig. 4 and Table 2 show examples of the deviation.

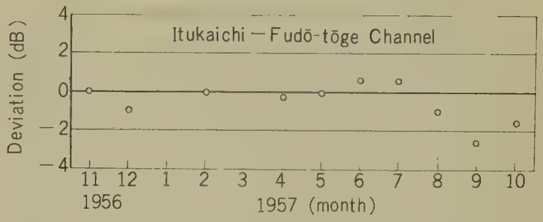


Fig. 4—Deviations of monthly median value from free space value. Path from Fudô-tôge to Itsukalchi.

The fading characteristics of the received power are very different according to the various conditions of the propagation path: for example, propagation distance, season, frequency, inclination of propagation path, and height of propagation path above the ground. So that a practical scale is necessary to express quantitatively the fading characteristics for various conditions of propagation.

In the case of a practical path, the wave reflected from the ground cannot be completely removed.

Therefore the residual fading due to interference by the ground-reflected wave must be taken into account. To ignore the quantity of the residual fading, we define the variation range (in dB) between 1 % of probability and 99 % of that determined by the cummulative distribution curve as the fading range, F_{rf}. And then it is made a statistical scale of the range of field variation which depends on the irregular distribution of the refractive index of the air for the given period

Table 2

Path length (km)	Frequency (Gc/s)	Month and year of number of months	Height (m)		Excess of path attenuation over free space loss for quoted percentage of time (dB)								
			Trans.	Rec.	99.9%	99.0%	90.0%	50%	10%	1%	0.5%	0.1%	0.01%
63	4	Sep., 1951	328	30	10.5	7.5	4	*a*	−4	−11.5	−13	−16	—
54	0.2	Dec., 1951	690	640	—	2.5	1.5	*a*	−1.5	−3.5	−4.5	—	—
	4				—	3	1.5	*a*	−2	−4	−4.5	−5	—
110	4	July., 1951	870	640	—	10	4	*a*	−7	−15	−17	—	—
45	4	Feb., 1952	70	50	—	6	2	*a*	−3.5	−6	−6.5	−10	—
63	4	May, 1952	1090	100	—	4	2	*a*	−3	−6	−7	−11	—
48	4	May, 1952	1090	600	—	3	1.5	*a*	−2	−4	−4.5	−5	—
150	0.2	Aug.~Sept., 1954	1090	600	—	6	2.5	*a*	−5.5	−15.5	−18	—	—
	1.5				—	5	3	*a*	−7	−20	−23	—	—
	4				—	9	5	*a*	−7	−22	−25	—	—
	6				—	12	6	*a*	−9	−23	−27	—	—
	1.5	June, 1954	1090	870	10	6	3	*a*	−5	−11.5	−12.5	—	—
	4				15	8	3	*a*	−4	−13.5	−15.5	−17.5	—
	6				10	7	3	*a*	−6	−17	−21	−31	—
100	1.5	Aug.~Sep., 1956	320	340	—	5.5	2.5	*a*	−4	−14	—	—	—
	4				—	9	3.5	*a*	−5	−14	—	—	—
	6.7				—	13	4.5	*a*	−5	−13.5	−17	—	—
	4	Nov., 1956	320	340	—	6	3	*a*	−3.5	−10	−12.5	—	—
	6.7				7.5	5.5	3.5	*a*	−4	−12.5	−14.5	—	—
	4	Aug., 1957	320	340	13	7	2.5	*a*	−7	−17.5	−21.5	—	—
	6.7				22	13	8	*a*	−9.5	−16	−18	—	—
50	1.5	July~Aug., 1956	350	50	—	5	3	*a*	−3.5	−10	−13.5	—	—
	4				—	5.5	3.5	*a*	−5.5	−12	−14	—	—
	6.7				—	7	4.5	*a*	−7	−13	−14	--16	—
44	1.5	Aug., 1958	240	40	4	3	2.5	*a*	−3	−7	−7.5	−10	—
	6.7				7	4.5	2.5	*a*	−3.5	−7.5	−8.5	−11	−16.5

Notes: Horizontal polarization, $a =$ within ± 2 dB.

over the propagation path.

The scale of fade-out from the median value, F_d, is given statistically.

$$F_d = 10 \log_{10} \frac{Z_{0.5}}{Z_{0.01}} (dB)$$

These scales are available for the discussion of the quality of communication circuits.

When the frequency distribution of the received power intensity (dB) is regarded as Gaussian distribution, the following relation between the standard deviation σ (dB) and F_{rf} (dB) obtains:

$$\sigma = \frac{F_{rf}}{4.66} \qquad (2)$$

If the distribution of received power intensity is assummed to be Gaussian and the median value of the power intensity is approximately estimated to be free space intensity, the distribution can be completely provided.

For many propagation tests, the frequency distribution of the value in dB of received power intensity does not always follow Gaussian distribution. It is often proper to approximate the frequency distribution of the power value (real value) by a Gamma distribution.

If it follows a Gamma distribution, the relation shown in Fig. 5 between F_{rf} (dB) and the index, ν, of Gamma distribution holds. Therefore the index is found by fitting the actually measured fading range F_{rf} to the theoretical range.

From many experimental results, the rough relation between F_d and F_{rf} is as follows:

$$\frac{F_d}{F_{rf}} = 0.65 \pm 0.06 \qquad (3)$$

As a result, if F_{rf} is found, F_d can be obtained.

1.4. Observed Characteristics of Fading

1.4.1. Long-Term Variation of Fading

For the path between Fudō-tōge and Itsukaichi, an example of the general aspect of the fading occurence during one year at 4 Gc (span length is about 100 km) is shown in Fig. 6.

Japanese standard time is shown as the abscissa, and the median value of hourly fading during a month is shown as the ordinate.

In this figure, quite remarkable diurnal characteristics are found. The cause is closely related to the diurnal variation[5] of the

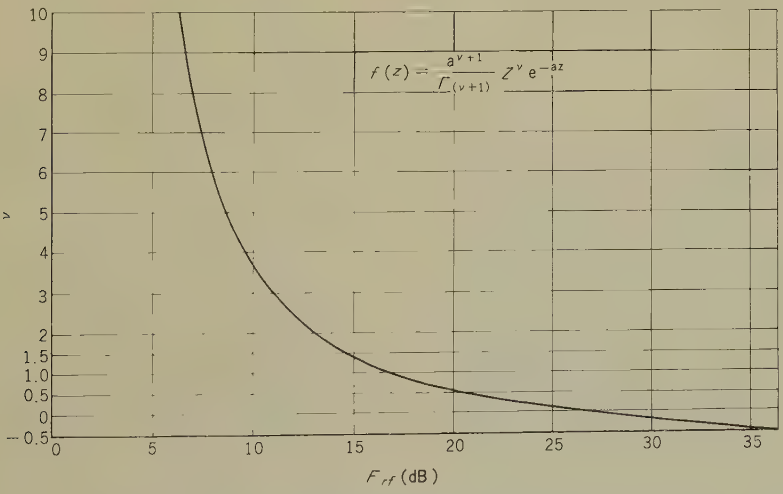


Fig. 5—Curve of ν against theoretical fading range F_{rf}.

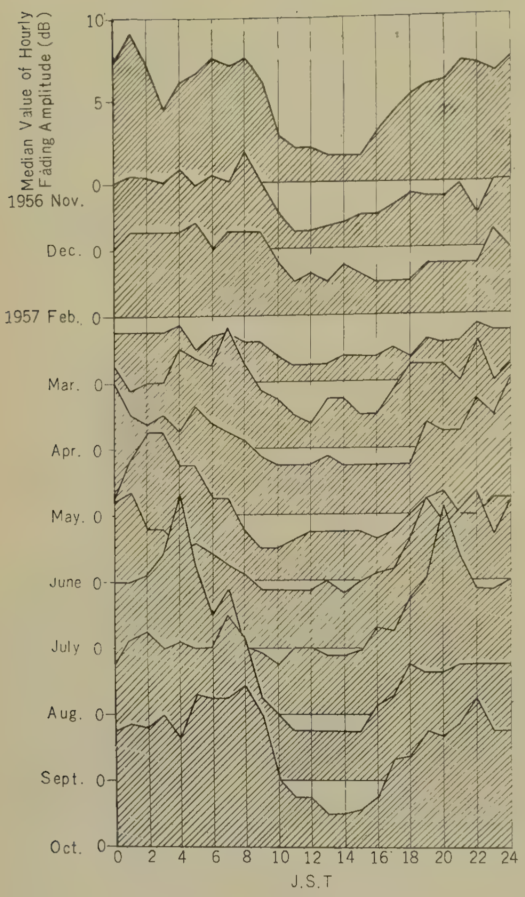


Fig. 6—Seasonal trend of median value of hourly fading amplitude. 6 path from Fudô-Tôge to Itsukaichi.

spatial refractive distribution.

For every month, the amplitude of fading is smaller in the daytime than at night. That is, the received power is more stable in the daytime than at night. Qualitatively, this characteristic is not so different from that due to the results of the fragmentary measurements made so far. In addition, the monthly variation due to the influence of season is clearly shown. As a whole, the variation is larger in summer than in winter. This tendency nearly corresponds to that supposed so far. As described later, according to fading estimations based on the correspondence of radiosonde data, the fading range in a normal year is smallest in winter, greatest in summer, and intermediate in spring and autumn. This is considered to be the typical form of the seasonal trend in a year with normal meteorological conditions. But many years are thought to have seasons with abnormal meteorological conditions.

So it is very dangerous to represent the general trend of seasonal variations by measured values for one season. The meteorological condition of that season must be compared with that of a normal year to verify whether there is anything abnormal or not.

Fig. 7 shows monthly variation of fading range (1–99 %) through a year. In Fig. 7 the meteorological parameter thought to have the closest correspondence is $U_\alpha/\langle\alpha\rangle$ and its monthly variation is also represented in the same figure, where $\langle\alpha\rangle$ shows the mean value of average gradient $\alpha=\Delta M/\Delta H$ (M.U./m) of M profile between the level of pressure 1000 mb and 900 mb obtained from radiosonde data. This data was taken at the Tateno Aerological Observatory near the mid-point of the propagation path. The value of U_α is the spuare root of the unbiased variance of

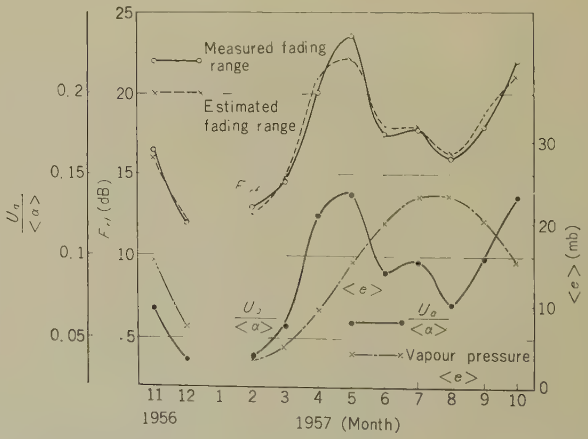


Fig. 7—Correlation among monthly fading range, $U_\alpha/\langle\alpha\rangle$ and vapor pressure $\langle e\rangle$.

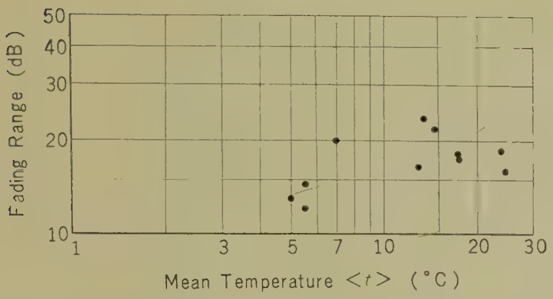


Fig. 8 (a)—Correlation between monthly fading range and monthly mean temperature.

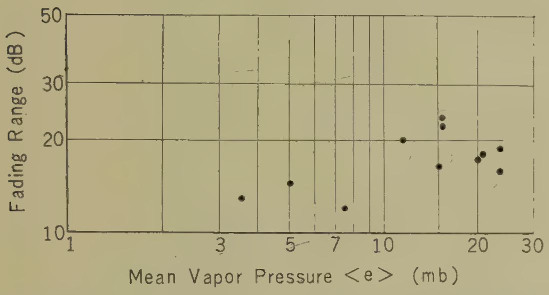


Fig. 8 (b)—Correlation between monthly fading range and monthly mean vapor pressure.

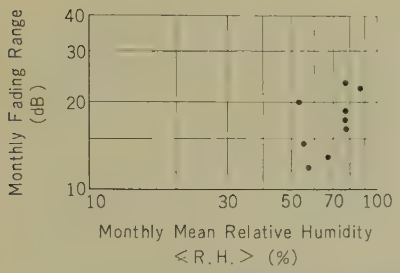


Fig. 8 (c)—Correlation between monthly fading range and monthly mean relative humidity.

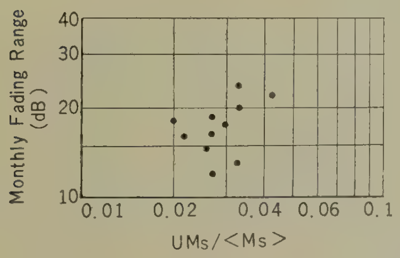


Fig. 8 (d)—Correlation between monthly fading range and monthly $U_{Ms}/\langle M_s\rangle$.

α during the period. The dotted line shows the estimated values obtained by using $U_\alpha/\langle\alpha\rangle$. The reason why such a good correspondence is obtainable will be described in detail after Section 2. In Fig. 8, its correspondence to other meteorological factors is discussed for a period covering a whole year. Fig. 8 (a), (b) and (c), show the correlation between monthly fading range and monthly mean value of temperature $\langle t\rangle$ on the ground, vapor pressure $\langle e\rangle$ on the ground, and relative humidity $\langle R.H\rangle$ on the ground. The (d) in the figure shows correlation between monthly fading range and $U_{M_s}/\langle M_s\rangle$.

The value of U_{M_s}, is the square root of the unbiased variance of M_s, and it is the modified refractive index on the ground. Here, special attention should be paid to the fact that the seasonal variation of fading has so far been explained by the vapor pressure and that the monthly mean value of vapor pressure on the ground is the radio-meteorological factor used for the estimation of fading range.

The seasonal variation of fading range, however, has often been found not to correspond to the variation of vapor pressure in recent years. The fading ranges for June～October shown in Fig. 7 are good examples. Accordingly, the mean value of vapor pressure on the ground is considered not enough for the explanation of the seasonal movement of fading variation, and besides not proper for the evaluation of fading range.

1.4.2. Mode of Fading

Examples shown in Fig. 9 (a) (b) are the typical shapes of fading record.

As shown by examples (a) and (b) of Fig. 9, there should be some qualitative correspondence, between the fading mechanism and fading shape. So it is considered not impossible to presume the fading mechanism to some extent from its shape.

In this case, from the fading shape itself, only few parts of the cause and mechanism are found.

Therefore it is necessary to find a method to determine its mechanism more accurately

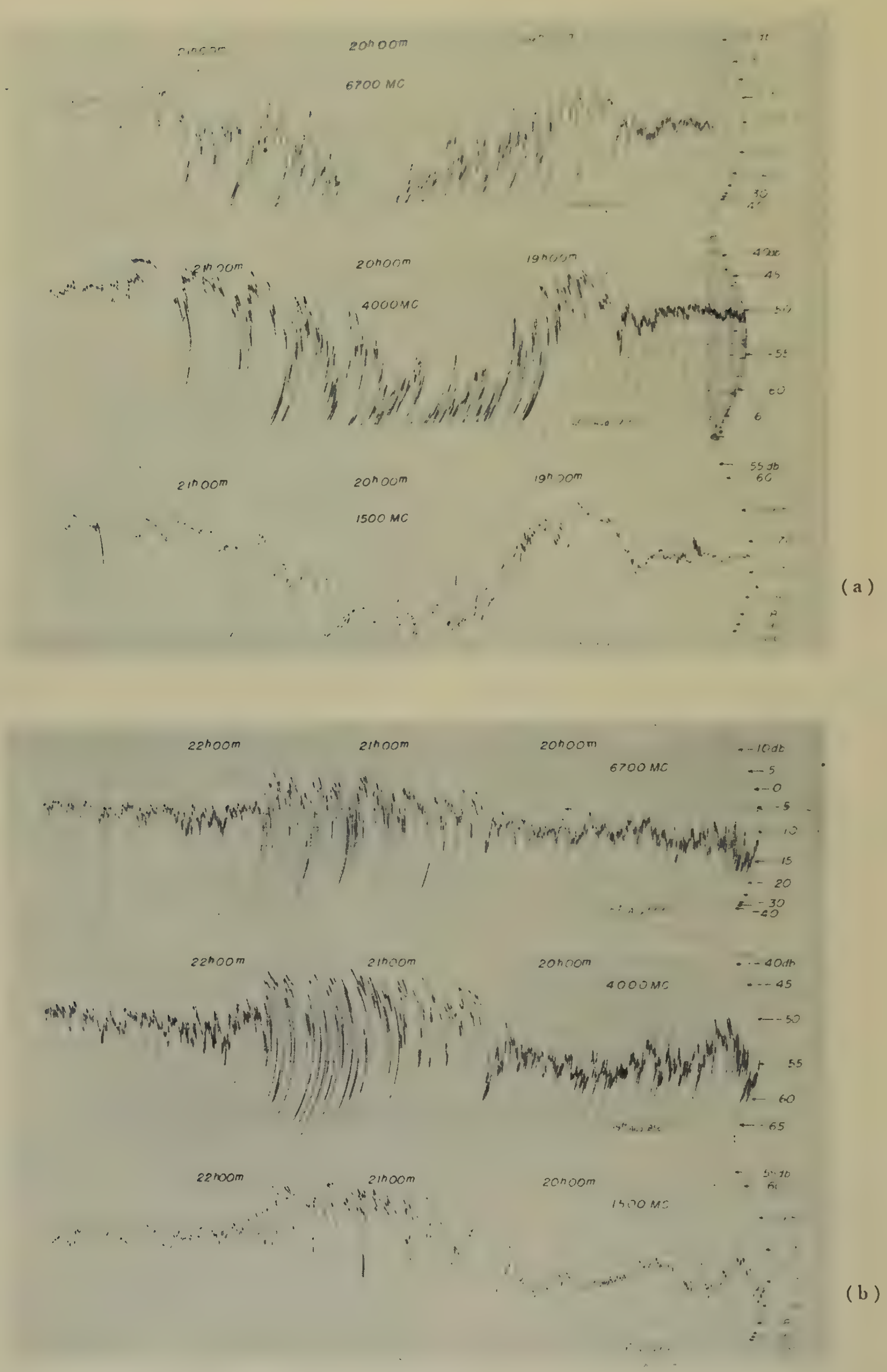


Fig. 9—Samples of simultaneous severe fading recorded on the path from Fudō-tōge to Itsukaichi at three frequencies 1.5 Gc/s, 4 Gc/s, and 6.7 Gc/s.

by obtaining more information. For this purpose, the frequency-diversity method was adopted.

Both Examples (a) and (b) are parts of simultaneous fading records of three frequencies (1.5 Gc, 4 Gc, 6.7 Gc) actually measured in the circuit between Fudō-toge, Mt. Tsukuba and Itsukaichi.

As shown in (a), at all frequencies, the average level of received power intensity begins to rise at about 1840 J.S.T., reaches the maximum level at about 1850 J.S.T. , begins to fade gradually, and then reaches the lowest level at about 2000 J.S.T.. Next it rises gradually until it comes back to the first level at about 2100 J.S.T.. In these records, the long-period variation, and the short-period variation which has more remarkable frequency dependence are observed.

In (b), long-period variation, which has little frequency characteristics and short-period variation which depends on frequency are also found. Fig. 9 shows examples of severe fading. On the contrary, Fig. 10 shows fading examples when fading is not so severe. The short-period variation superposed is smaller in every instance than the variation in Fig. 9.

What is noteworthy in these examples, is that the fading variation when fading is not so severe does not have remarkable frequency dependence. The qualitative characteristics above are found in common to nearly all fading records of propagation tests. They are summed up as follows:

1) There are two sorts of fading—long-

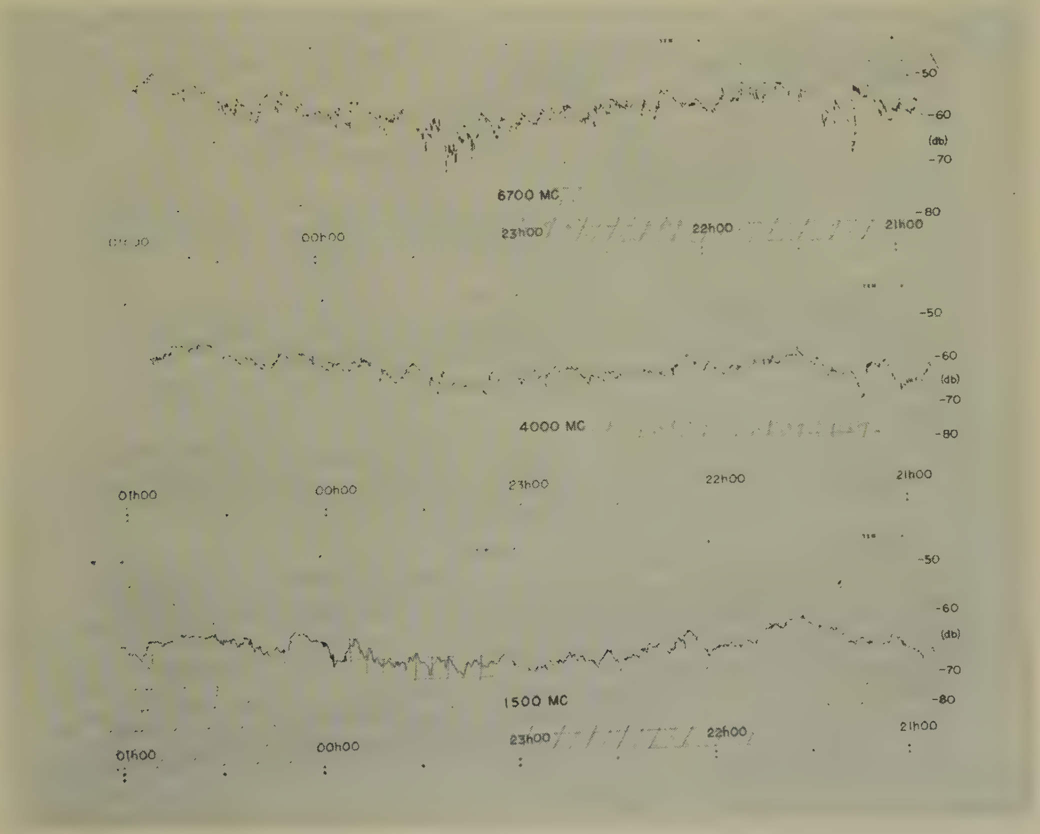


Fig. 10—Samples of simultaneous light fading recorded on the path from Fudō-tōge to Itsukaichi at three frequencies 1.5 Gc/s, 4 Gc/s, and 6.7 Gc/s.

period variation and short-period variation. They show various fading modes because they are often superposed on each other.

2) Usually, the long-period of variation does not vary much with frequency.
3) The short-period variation mode varies significantly depending on frequency. The higher the frequency is, the shorter the variation period is and the larger the variation amplitude is.
4) It is thought that there is some difference between the causes of the long-period variation and the short-period one, or between the mechanisms of both the variations.

1.4.3. Long-Period Variation

If the fading records of propagation tests using simultaneously two or more frequencies are observed in detail paying attention to the variation modes mentioned above, then a sort of average level of received power is imagined which varies commonly with time and almost independently of frequency.

An attempt has been made to show this assumed average level in Figs. 11 and 12.

The points plotted in Fig. 11 show the signal level of the interfered wave which is calculated from the maximum fading amplitude during 5 minutes on the assumption that the fading mechanism is the interference of two rays propagating through the irregular atmosphere and that the phase difference between the two rays changes widely with time. This example is dealt with for the path between Mt. Tsukuba and Mt. Futago.

The levels shown respectively in Fig. 12 (a), (b), and (c) are envelopes of maximum value of received power. Here, it is approximately assumed that the number of multiple rays does not depend on frequency and that the phase of the multiple rays changes widely.

Of course, the levels embodied by these two methods do not completely correspond to

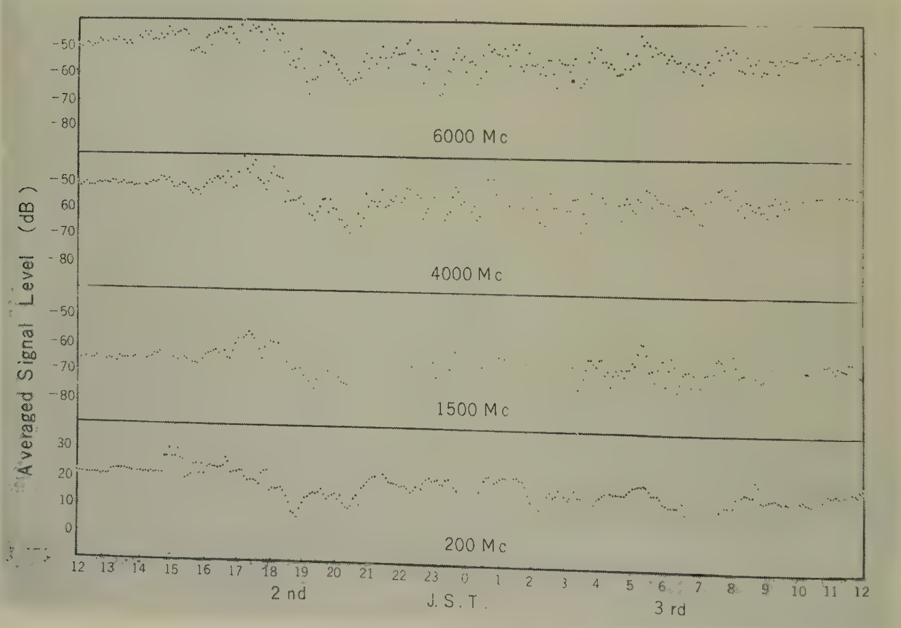


Fig. 11—A sort of averaged level calculated under the condition that fading is due to the interference of two rate. At four frequencies 200 Mc/s, 1500 Mc/s, 4000 Mc/s, and 6000 Mc/s on the path from Mt. Futago to Mt. Tsukuba.

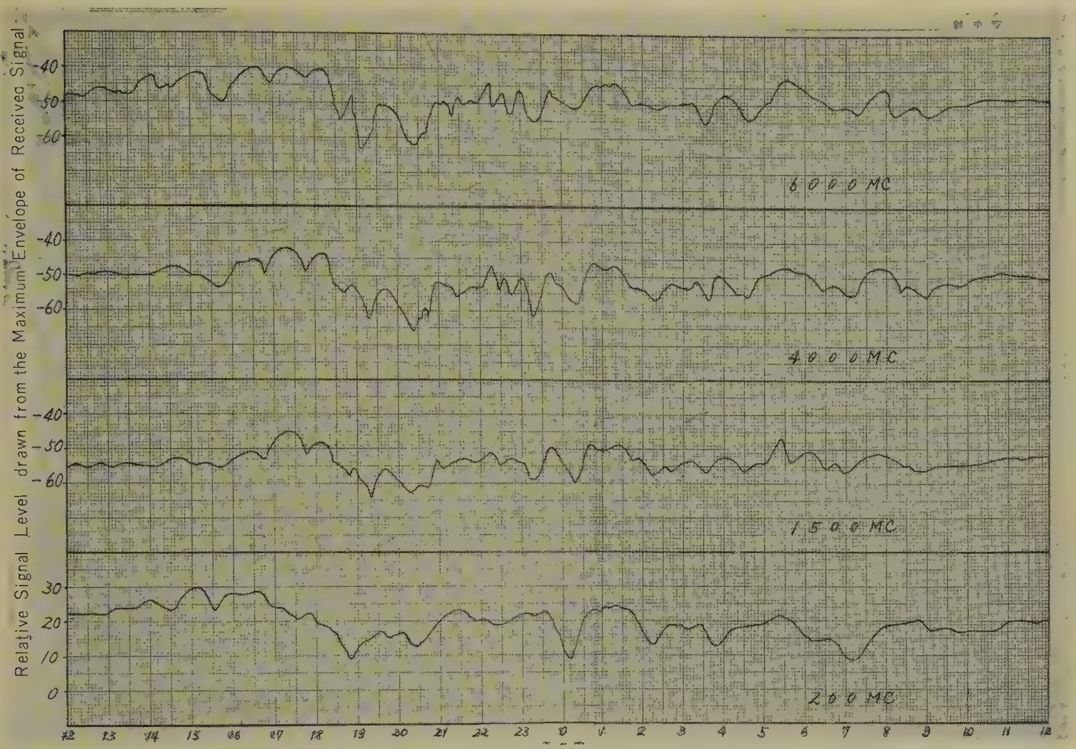


Fig. 12 (a)—Examples of long period signal variation levels taken from the fading records received on the path from Mt. Futago to Mt. Tsukuba at four frequencies 200 Mc/s, 1500 Mc/s, 4000 Mc/s, and 6000 Mc/s. (The scales on ordinate and abscissa indicate respectively dB and J.S.T..)

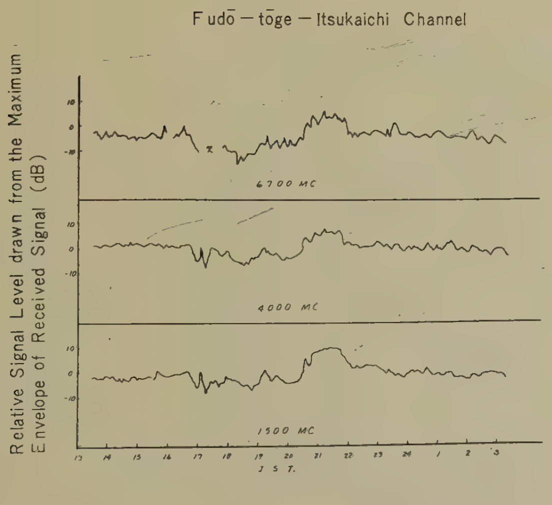


Fig. 12 (b)—Examples of long period signal variation levels taken from the fading records received on the path from Fudō-tōge to Itsukaichi at three frequencies 6700 Mc/s, 4000 Mc/s, and 1500 Mc/s.

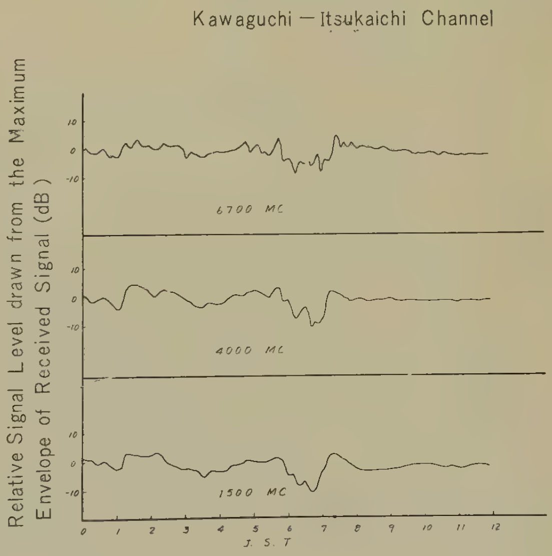


Fig. 12 (c)—Examples of long period signal variation levels taken from the fading records received on the path from Kawahuchi to Itsukaichi at three frequencies 6700 Mc, 4000 Mc, and 1500 Mc/s.

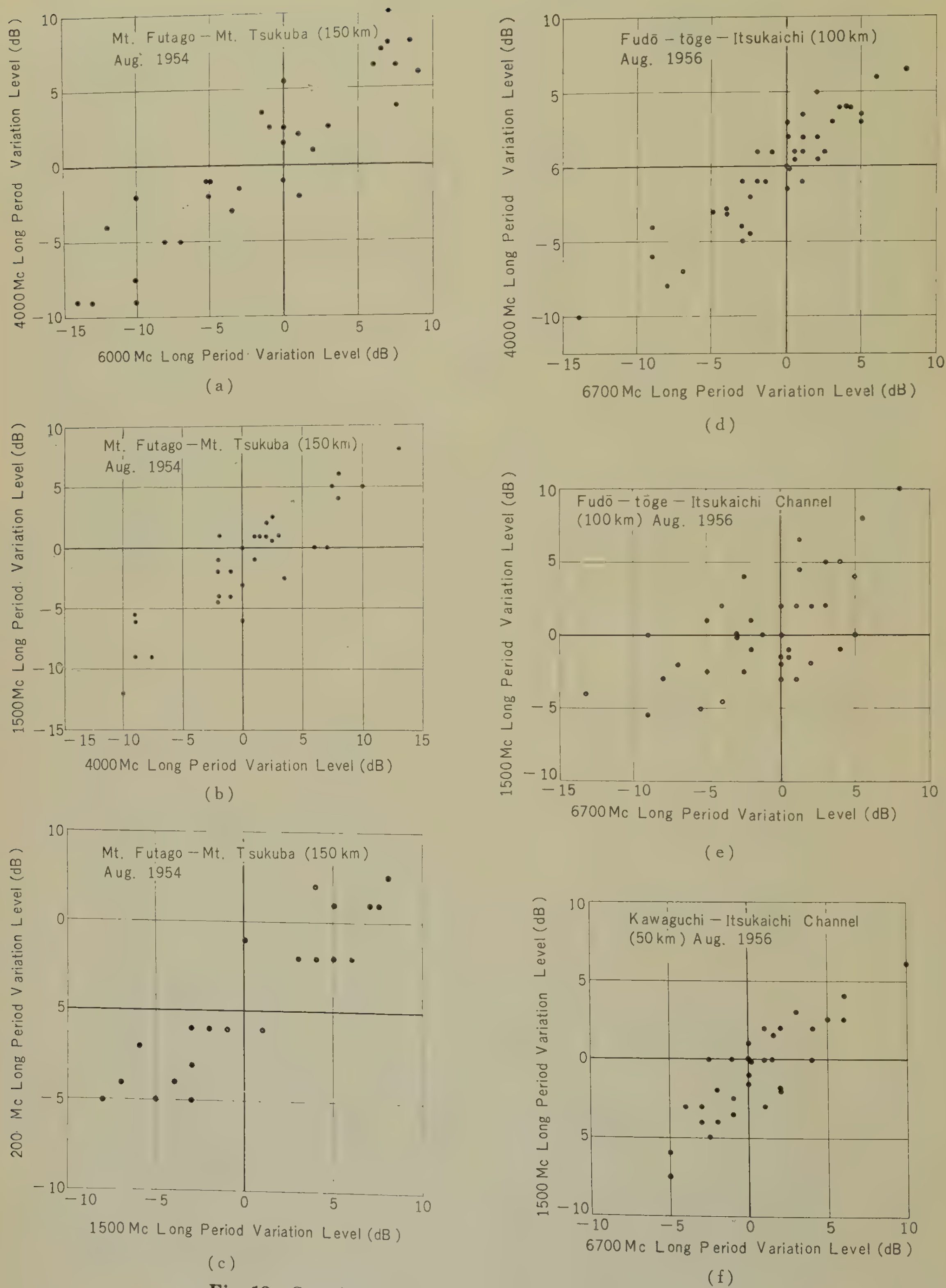


Fig. 13—Correlation between long period variation levels.

each other, but the macroscopic relative tendencies of the level variation show a fairly good correspondence to each other. As the levels expressed in Figs. 11 and 12 have been obtained by the above method, the frequency effect has not been completely removed. Therefore, they are not exactly the abstracted long-period level variation.

But macroscopically, the level movements can be said to correspond fairly well to each other.

Examples of the correlation between these level movements are shown again in Fig. 13 (a)–(f). From these figures, the levels based on the above method, are thought nearly the same as the long-period variation levels assumed by us.

1.4.4. Short-Period Variation

It is as difficult to draw short-period variation as to draw long-period variation from fading record concretely. In General, variation with time are mainly characterized by their amplitude and period. Probably there are various ways to express variation amplitude, but now, as mentioned above, the maximum variation amplitude (dB) of instantaneous received power during 5 minutes will be taken as the sample, and then the correlation between the variation amplitude and period will be obtained.

Fig. 14 shows a comparison among the variation amplitudes for the frequencies of 1.5 Gc, 4 Gc, and 6 Gc, for severe fading produced in the propagation path between Mt. Futago and Mt. Tsukuba. The correlation between variation amplitudes for all frequencies is very remarkable.

That is, the higher frequency becomes, the larger the variation amplitude grows. Moreover, it is found that the variation very severely depends on frequency; also that the time of advent and of extinction for all frequencies correspond very well to one another. By closer observation, it is found that at a higher frequency the variation appears earlier and disappears later; i. e., the variation continues longer.

Fig. 15 shows the relation of fading amplitude vs. frequency where 4 Gc is taken as the standard frequency.

Subsequently, consideration will be made on the period of variation. The period is

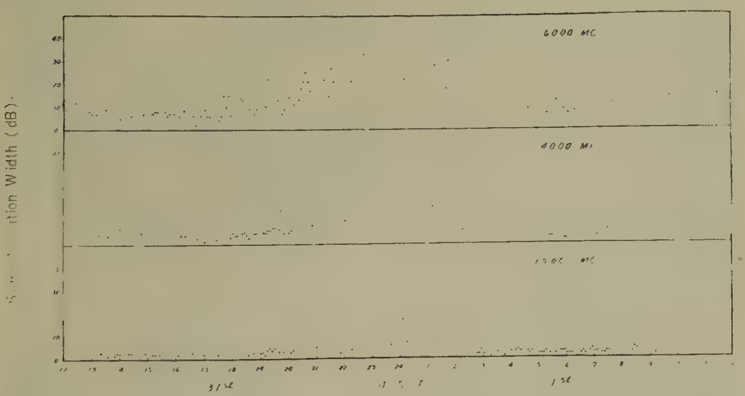


Fig. 14—Frequency dependence on fading width received on the path from Mt. Futago to Mt. Tsukuba at three frequencies 1500Mc/s, 4000 Mc/s, and 6000 Mc/s.

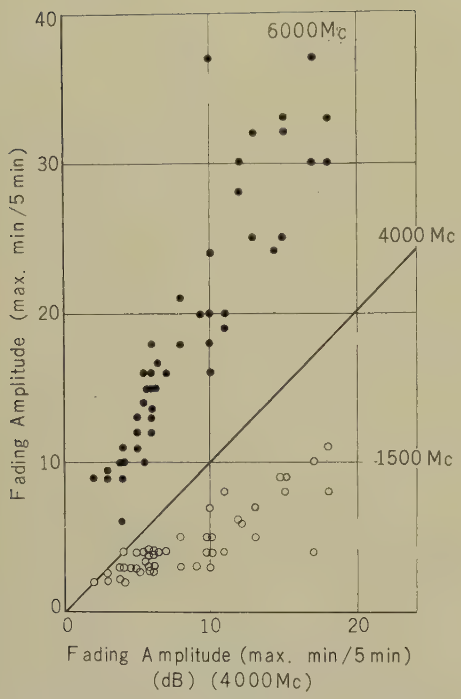


Fig. 15—Relation of fading amplitude against frequency.

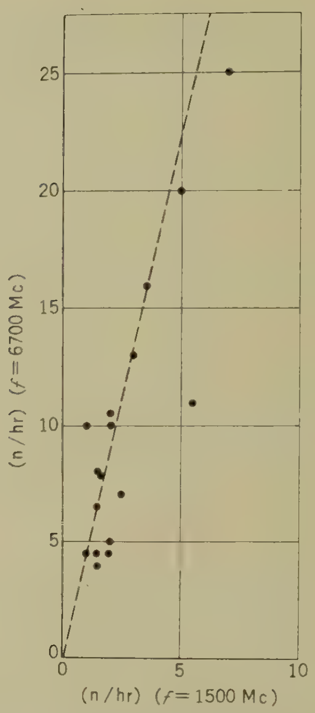


Fig. 16 (b)—Correlation of fading rate between two frequencies 6700 Mc and 1500 Mc.

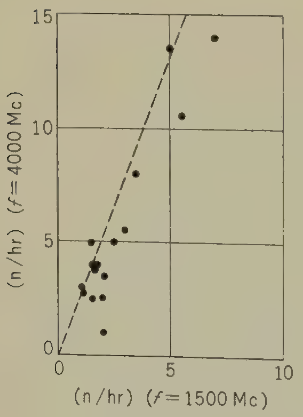


Fig. 16 (a)—Correlation of fading rate between two frequencies 4000 Mc and 1500 Mc.

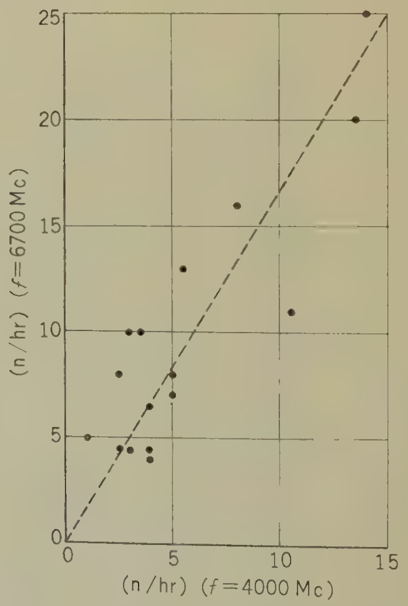


Fig. 16 (c)—Correlation of fading rate between two frequencies 6700 Mc and 4000 Mc.

uaually expressed by fading rate.

As an example, the 16 hours when severe fading appeared in both the propagation paths between Fudō-tōge and Itsukaichi, and between Kawaguchi and Itsukaichi were selected.

The frequency correlation between the fading rates (n/hr) of the waves at 1.5 Gc, 4 Gc, and 6.7 Gc is shown in Fig. 16 (a)-(c).

In the figure, the dotted-line is drawn by assuming that n is proportional to frequency.

As a result of calculation of the correlation coefficient r, the following relations are obtained:

$$r_{1.5,\,4.02}=0.661$$

$$r_{4.02,\,6.7}=0.745$$

$$r_{1.5,\,6.7}=0.847$$

Considering that the number of samples is limited, the significant examination of the correlation was made by R. A. Fisher's theory. As a result of the examination, the correlation coefficient r can be thought significant.

Concerning each correlation coefficient, values of regression coefficient, $m(f_i/f_j)$ of ordinate value to abscissa value were obtained as follows:

Table 3

RELATION BETWEEN FADING RATE AND FREQUENCY RATIO

f_i (Gc)	f_j (Gc)	f_i/f_j	$m(f_i/f_j)$
6.7	1.5	4.47	3.88
4.02	1.5	2.66	2.58
6.7	4.02	1.67	1.73

As a result, it is found that the fading rate is nearly proportional to the radio wave frequency.

This is the result supposed from the theory of interference between two waves. When minute variation characteristics are discussed, various characteristics of the short-period variation, such as the statistics of the frequency of fade-out below a specially fixed level of received power, and the time length of fade-out, must be clarified.

As an example, Fig. 17 shows the frequency distribution of fade-out time below the −10 dB and −5 dB levels at 4 Gc/s and 6 Gc/s (Mt. Futago–Mt. Tsukuba propagation path) respectively, when the free space received power intensity is taken as the standard level. The higher the frequency of the radio wave is, the lower the fade-out level is and the more frequently short-term fading occurs.

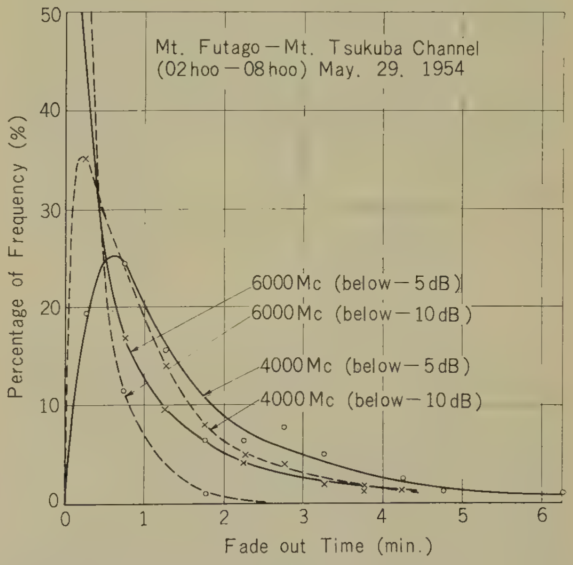


Fig. 17—Example of frequency distribution of fade-out time below a certain level on the path from Mt. Futago to Mt. Tukuba (02h00～08h00 May, 29, 1954).

Fig. 18 is an example of the minute structure of the fading shape. It shows shapes of fading at 3.970 Gc, 4.02 Gc, and 4.15 Gc measured at the same time by pen-writing oscillographs, and shows that the time that they fade out below the given reference level depends on frequency. Moreover, it is found that the fading depth and the outage time have a little relation to frequency.

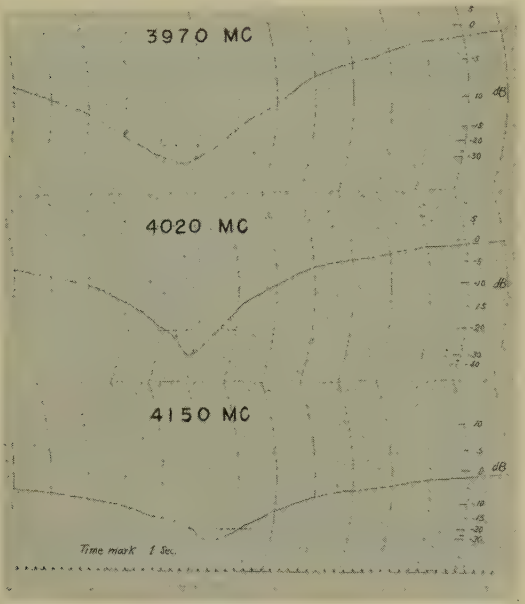


Fig. 18—Example of fine structure of fading received on the path from Asahi to Toyama.

These characteristics are also found in other propagation paths. Synthesizing all the experimental results together, therefore, the main cause of short-period variation is considered to be multiple-ray interference. This fact plays a very important role in communication, and indicates that as a countermeasure to the variation of received field, the diversity method is most effectively free from short-period variation.

2. Relation between Fading and the Variation of the Refractive Index of the Air

2.1. Qualitative Consideration of Fading Mechanism

For the purpose of physical explanation of these two characteristics of long- and short-period fading classified phenomenally, not only experimental analysis of fading but also theoretical analysis based on measurements of spatial distribution of the structure of the refractive index of the air is necessary.

Necessary and precise data cannot be easily obtained by means of the refractive index measuring apparatus (temperature, humidity and pressure) now in practical use, because its sensitivity and accuracy are insufficient, and the measurement is too troublesome. There is much left to improve this apparatus and to develop new techniques of measuring fading.

In general, phenomenally, long-period level variation occurs with fading and also short-period variation superposed on it grows more remarkable as the fading becomes more severe. How much is it superposed on?

It is too complicated to find the tendency. The reason is thought to be that the cause of long-period variation is not always that of short-period variation.

To explain the variation of long-period level, the above stated mechanism which is almost independent on frequency must be assumed. On the other hand, as is found from the fading characteristics shown already, the short-period variation has such remarkable frequency effects (the period and amplitude of variation), that the interference of plural waves must be mainly assumed as the mechanism.

Now, if the model of the M profile shown in Fig. 19 is assumed, it can be explained by geometrical optics that the path of the rays is changed by the bending part of the M profile, and that as a result the spatial energy density at the receiving point can vary.[1) 2) 4)] The change of this energy density will be named convergence effect or divergence effect.

The long-period variation is considered to be due to the macroscopic irregularity of the spatial distribution of the refractive index of the air.

As shown in Fig. 19, the short-period variation is caused when the atmosphere is constructed in such a way as to produce multiple propagation paths. This sort of variation

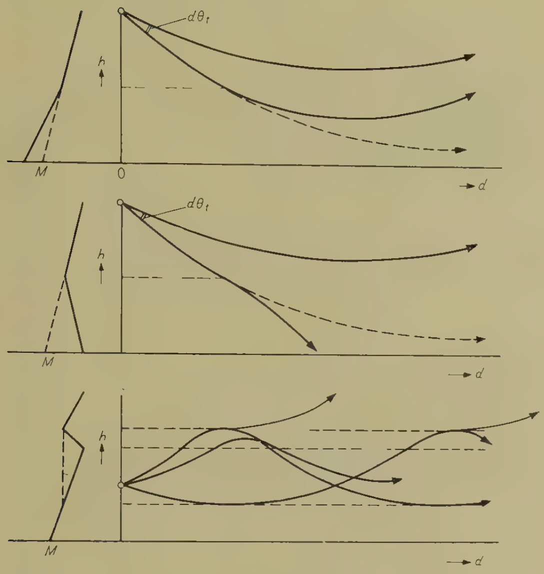


Fig. 19—Model of modified index profile and path of ray.

is considered to be due not only to the above-mentioned macroscopic irregularity but also due to the non-uniformity of the atmosphere on a small scale. When the change of gradient occurs partially in the gradient of the refractive index, it is expected that long-period variation is caused in the process of its appearance or extinction. If parts of the gradient develop into radio ducts, multiple-propagation paths will be produced at the same time by the duct.

Qualitatively, it is not very unreasonable to think that in general the long- and short-period variations simultaneously appear due to the complicated microscopic irregularity of the atmosphere. Fading can be explained qualitatively quite well by such considerations.

2.2.1. M Profile and Field Intensity

The field intensity in a stratified atmosphere with nonlinear M Profile $M(h)$ was calculated theoretically by W. H. Furry,[4] C. L. Pekeris,[4] etc.[2] and Goto.[5] According to Goto's formula, the received field intensity E is given as:

$$E \approx \frac{3}{2} \frac{\sqrt{10P\lambda}}{\sqrt{D}} \sum_{s=0}^{\infty} \varphi_s(h_1)\varphi_s(h_2)$$

$$\times \exp jm_s \frac{2\pi}{\lambda} D \Big| \quad (\text{Volt/m}) \qquad (4)$$

where P: Transmitted power (Vertical polarized wave) (Watt) (w)
E: Received field intensity (Volt/Meter) (V/m)
λ: Wave length (Meter) (m)
$h_1(h_2)$: Transmitting (Receiving) antenna height (Meter) (m)
D: Distance (Meter) (m)

$\varphi_s(h_1), \varphi_s(h_2)$ are height gain functions, and are obtained from the next equation:

$$\frac{d^2\varphi_s(h)}{dh^2} + \frac{8\pi^2}{\lambda^2} 10^{-6}\{M(h) - M_s\}\varphi_s(h) = 0$$

When $h \to \infty$, they show travelling waves. They are such normalized functions as tend nearly to zero with $h \to 0$ (on the surface).

$m_s (= M_s \cdot 10^{-6})$ is the root of the following equation:

$$\frac{d}{dh}\varphi_s(o) + j\frac{2\pi}{\lambda}\left\{\frac{\sqrt{\kappa_e + j\,60\sigma\lambda - 1}}{\kappa_e + j\,60\sigma\lambda}\right\}\varphi_s(o) = 0$$

where κ_e: Inductivity of the earth.
σ: Conductivity of the earth.

Accordingly, if the values of the M profile $M(h)$ are accurately measured in succession, and the field intensity is calculated by substituting the measured M value into Eq. (4), the statistics of the nonlinear M profile should be theoretically related to those of fading caused by the M profile.

Within the visible region, the series in Eq. (4) does not converge very well.

Now, it is not easy to measure accurately $M(h)$ of various propagation paths, because the accuracy of the equipment now in use is not satisfactory. So another method will be experimentally obtained by measured correlation between the M profile and field intensity.

2.2.2. Relation between the Statistics[6)] of the Structure of the Distribution of the Refractive Index of the Air and Fading

As mentioned above, fading is phenomenally considered to be constructed by the superposition of long- and short-period variation.

The short-period variation is considered due to the interference of multiple-rays. But the general characteristics of multiple rays; such as, the difference of path length among multiple rays, the number of the rays, and their intensities, are not sufficiently well understood.

And so it is thought that the estimation of short period fading range from the fading records is, at present, the possible method.

As will be described in detail in Section 3, if fading can be expressed by the product of the instantaneous value of the long-period variation level and that of short-period one, the quantity of fading can be expressed by the combination of the fading quantities of long- and short-period variation. Paying attention to this idea, the fading quantity of long-period variation will be considered.

To simplify the problem, it is assumed that the modified refractive index M of the air varies only with height.

From the propagation path shown in Fig. 20, suppose that the angles of beam radiated at transmitting point h_t are shown by θ_t, $\theta_t+d\theta_t$ and ϕ_t, $\phi_t+d\phi_t$ respectively and that the heights where the beams reach the distance S from the transmitting point are h_r and h_r-dh_r respectively.

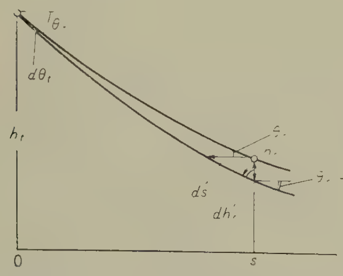


Fig. 20—Geometry of path of ray.

Also let θ_r and $\theta_r+d\theta_r$ be the arrival angles at the receiving point.

Considering the directivity $g_t(\theta_t, \phi_t)$ of the transmitting antenna, the density per unit area at distance S radiated within the small radiation angles of $d\theta$ and $d\phi$ becomes:

$$g_t(\theta_t, \phi_t)\cdot d\theta_t\cdot d\phi_t/S\, d\phi_t(-dh_r)\cos\theta_r$$

$$=g_t(\theta_t, \phi_t)/S\sin\theta_r\cdot\frac{(-dS)}{d\theta_t}$$

If x is defined as the ratio of the energy density in the atmosphere to the density in the idealized standard atmosphere, the energy density x is given by

$$x=\frac{\dfrac{g_t(\theta_t, \phi_t)d\theta_t\cdot d\phi_t}{S^2 d\theta_t\cdot d\phi_t}}{\dfrac{g_t(\theta_t, \phi_t)}{S\sin\theta_r\cdot\dfrac{(-dS)}{d\theta_t}}}$$

$$=\frac{(-dS)\sin\theta_r}{S\,d\theta_t}\doteqdot\frac{\theta_r(-dS)}{S\,d\theta_t}.$$

$$S=\int_{h_r}^{h_t}\theta^{-1}dh$$

where $\theta=\sqrt{\theta_t^2+2(M-M_t)}$

M_t : The value of M at the transmitting point h_t

M : The value of M at a given height h

x is rewritten as

$$x=\frac{\theta_t\theta_r}{S}\int_{h_r}^{h_t}\theta^{-3}dh. \tag{5}$$

At first, suppose that the M profile is linear and varies only with time.

In this case, if

$$M=M_t-\alpha(h_t-h) \tag{6}$$

is put into Eq. (5), the following relation is obtained.

$$x=\frac{\theta_t\theta_r}{S}\int_{h_r}^{h_t}\theta^{-3}dh=\frac{\theta_t\theta_r}{\alpha S}\int_{\theta_r}^{\theta_t}\theta^{-2}\,d\theta=1.$$

where $\alpha \equiv \frac{dM}{dh}$.

The above equation shows that in an atmosphere of linear M, even if the gradient changes with time, energy density never changes.

Next, suppose the M profile is not linear. In this case, M can be given by the function of height h and time t. In Eq. (6), suppose

$$M = \langle M \rangle + \varDelta M \quad \text{and}$$

$$M_t = \langle M_t \rangle + \varDelta M_t,$$

where $\langle \quad \rangle$ shows the mean value for time, and $\varDelta$ shows the deviation from the mean value at any time.

If x is taken in dB,

$$x = 10\ \mathrm{long}_{10}\, x$$

then the following relation is obtained.

$$X(\mathrm{dB}) =$$

$$m\left\{\log_e \theta_t + \log_e \theta_r + \log_e\left(\int_{h_r}^{h_t} \theta^{-3}\, dh\right) - \log_e S\right\} \tag{7}$$

where $m = 10 \log_{10} e$.

Let the variation of X be δX, and if the second order and the higher orders of its variation are assummed to be small and neglected, then

$$\delta X = m\left\{\frac{\delta\theta_t}{\theta_t} + \frac{\delta\theta_r}{\theta_r} + \frac{\delta\int_{h_r}^{h_t} \theta^{-3}\, dh}{\int_{h_r}^{h_t} \theta^{-3}\, dh} - \frac{\delta S}{S}\right\}. \tag{8}$$

Under the condition that S is constant, the following equations are obtained.

$$\delta S = \delta_{\theta_t} \cdot S + \delta_M S$$

$$= -\theta_t \int_{h_r}^{h_t} \theta^{-3} dh\, \delta\theta_t - \int_{h_r}^{h_t} \theta^{-3} \delta M dh = 0,$$

$$\delta\theta_t = -\frac{\int_{h_r}^{h_t} \theta^{-3}\, \delta M dh}{\theta_t \int_{h_r}^{h_t} \theta^{-3}\, dh},$$

$$\delta\theta_r = \delta_{\theta_t} \theta_r + \delta_M \theta_r = \frac{\theta_t \delta\theta_t + \delta M_r}{\theta_r},$$

$$\delta \int_{h_r}^{h_t} \theta^{-3}\, dh = \delta_{\theta_t} \int_{h_r}^{h_t} \theta^{-3}\, dh + \delta_M \int_{h_r}^{h_t} \theta^{-3}\, dh$$

$$= -3\left\{\theta_t \int_{h_r}^{h_t} \theta^{-5}\, dh\, \delta\theta_t + \int_{h_r}^{h_t} \theta^{-5}\, \delta M dh\right\}.$$

If these equations are substituted into Eq. (8) on the assumption that the mean shape of the M profile with respect to time is linear, then the following equation is obtained.

$$\left.\begin{aligned} &\delta X = m \int_{h_r}^{h_t} w \delta\alpha dh. \\ &\text{where} \\ &w = \frac{1}{\theta_t - \theta_r}\left(1 - \frac{\theta_t \theta_r}{\theta^2}\right)\frac{1}{\theta} \\ &\delta\alpha \equiv \delta\frac{dM}{dh} = \frac{d}{dh}(\delta M) \end{aligned}\right\} \tag{9}$$

If δM is proportional to h, then $\delta X = 0$ is obtained.

As shown in Reference (6), the M profile of the atmosphere is complicated. According to the theory in Reference (6), the structure of the atmosphere can be expressed as the following correlation function.

Let $\rho_\alpha(\varDelta h)$ be the correlation function of α against height

$$\rho_\alpha(\varDelta h) = \begin{cases} 1 & 0 \leq \varDelta h \leq \varDelta h_0 \\ 0 & \varDelta h > \varDelta h_0 \end{cases} \tag{10}$$

And suppose that this relation holds approximately in the propagation path and that the atmosphere is stratified consisting of layers with thickness $\varDelta h$ and gradient α_i which are randomly sampled from the assembly of α_0.

Then, Eq. (9) can be written as:

$$\delta X = m \sum_{i=1}^{n} (\delta\alpha)_i \int^{(\Delta h)_i} w\,dh \tag{11}$$

If the mean square deviation of the both sides of Eq. (11) are taken,

$$\langle \delta X^2 \rangle \equiv \sigma_X{}^2, \qquad \langle \delta\alpha^2 \rangle \equiv \sigma_\alpha{}^2,$$

then, in general,

$$\sigma_X{}^2 = m^2 \Bigg\{ \sum_{i=1}^{n} (\sigma_\alpha{}^2)_i \int \left(\int^{(\Delta h)_i} w\,dh \right)^2$$

$$+2 \sum_{j \neq k} \rho_{j,k} (\sigma_\alpha)_j (\sigma_\alpha)_k$$

$$\times \left(\int^{(\Delta h)_j} w\,dh \right) \left(\int^{(\Delta h)_k} w\,dh \right) \Bigg\}. \tag{12}$$

where $\rho_{j,k}$ is the correlation coefficient between $(\sigma_\alpha)_j$ and $(\sigma_\alpha)_k$.
If we assume, $\rho_{j,k} = 0 \quad (j \neq k)$,

$$(\sigma_\alpha)_i = \sigma_{\alpha 0}, \qquad \langle \alpha_0 \rangle_i = \langle \alpha_0 \rangle = \langle \alpha \rangle,$$

$$\sigma_X{}^2 = m^2 \sum_{i=1}^{n} \sigma_{\alpha 0}{}^2 \left(\int^{(\Delta h_0)_i} w\,dh \right)^2.$$

And then it is rewritten as:

$$\sigma_X{}^2 \approx m^2 \sigma_{\alpha 0}{}^2 \cdot \Delta h_0 \int_{h_r}^{h_t} w^2\,dh \tag{13}$$

According to experiments, this assumption is nearly satisfied.[6)]

i) In case $\theta_r \neq 0$
If w of Eq. (9) is substituted into Eq. (13),

$$\sigma_X{}^2 \doteqdot m^2 \frac{\sigma_{\alpha 0}{}^2 \cdot \Delta h_0}{\langle \alpha \rangle^2} \cdot \frac{4}{3} \cdot \frac{u^2}{\Delta H} \quad (u \ll 1) \tag{14}$$

where $u = \dfrac{\langle \alpha \rangle d^2}{2\Delta H}$, $\quad \Delta H = h_t \sim h_r$

and d is the distance between the transmitting and receiving points

ii) In case $\theta_r \to 0$
If Eq. (13) is rewritten as

$$\sigma_X{}^2 = m^2 \sigma_{\alpha 0}{}^2 \cdot \Delta h_0 \left\{ \frac{1}{\Delta h_0} \left(\int_{h_r}^{h_r + \Delta h_0} w\,dh \right)^2 \right.$$

$$\left. + \int_{h_r + \Delta h_0}^{h_t} w^2\,dh \right\}.$$

If $\theta_r \neq 0$, from Eq. (9)

$$\int_{h_r}^{h_t} w\,dh = 0.$$

Therefore it can be considered that

$$\lim_{\theta_r \to 0} \int_{h_r}^{h_t} w\,dh \to 0, \quad \text{if} \quad \lim_{\theta_r \to 0} \int_{h_r}^{h_t} w\,dh \ \text{exists.}$$

Accordingly,

$$\left(\lim_{\theta_r \to 0} \int_{h_r}^{h_t + \Delta h_0} w\,dh \right)^2$$

$$= \left(-\lim_{\theta_r \to 0} \int_{h_r + \Delta h_0}^{h_t} w\,dh \right)^2 = \frac{(\theta_t - \Delta\theta)^2}{\langle \alpha \rangle^2 \theta_t{}^2}$$

As a result,

$$\sigma_X{}^2 = m^2 \frac{\sigma_{\alpha 0}{}^2}{\langle \alpha \rangle^2} \left\{ \left(1 - \frac{\Delta\theta}{\theta_t} \right)^2 + \frac{\Delta\theta^2}{2\theta_t{}^2} \log_e \frac{\theta_t}{\Delta\theta} \right\}$$

or

$$\sigma_X{}^2 = m^2 \frac{\sigma_{\alpha 0}{}^2}{\langle \alpha \rangle^2} \left\{ \left(1 - \frac{\Delta d}{d} \right)^2 \right.$$

$$\left. + \frac{1}{2} \left(\frac{\Delta d}{d} \right)^2 \log_e \frac{d}{\Delta d} \right\},$$

$$\sigma_X{}^2 = m^2 \frac{\sigma_{\alpha 0}{}^2 \cdot \Delta h_0}{\langle \alpha \rangle^2} \left\{ \left(\frac{1}{\sqrt{\Delta h_0}} - \frac{1}{\sqrt{\Delta H}} \right)^2 \right.$$

$$\left. + \frac{1}{4\Delta H} \log_e \frac{\Delta H}{\Delta h_0} \right\} \tag{15}$$

where

$$\Delta\theta = \sqrt{2\langle \alpha \rangle \cdot \Delta h_0},$$

$$\Delta d = \sqrt{\frac{2\Delta h_0}{\langle \alpha \rangle}},$$

$$d=\sqrt{\frac{2\Delta H}{\langle\alpha\rangle}},\qquad \Delta H=h_t\sim h_r$$

By means of the above approximate calculation it is found that the long-period variation is related to $\sigma_{\alpha 0}\cdot\sqrt{\Delta h_0}$ and that when θ_r becomes small there is a tendency for σ_X to increase rapidly. An experimental analysis of fading range will be made, paying attention to these characteristics.

3. General Characteristics and Method of Quantitative Estimation of Fading

3.1. Structure of Fading

As already discussed, it is difficult to select accurately and concretely only the long-period variation level from the fading records.

However difficult it may be, it is thought possible to imagine approximately the variation level having nothing to do with frequency.

Now, this variation level is defined anew to be long-period variation level, and the instantaneous value of this level is represented by x.

According to investigations of many fading records, it is assumed that the instantaneous received power intensity z is expressed by the product of both the instantaneous level x of long-period variation and that y of short period variation. On this assumption, if the distribution function to which x and y belong and the correlation between x and y are found, the distribution which the instantaneous value of received power intensity z $(=x\cdot y)$ follows can be calculated, and as a result, fading range can be represented by the statistics of x and y. But, it is almost impossible to obtain exact correlation between x and y from the fading record. Accordingly, research will be made on the tendency of correlation, concerning the approximate values of x and y.

The values in dB of x, y, and z are shown respectively by X, Y, and Z. Assume that the quantity corresponding to x is X (dB), the instantaneous value of mean received power level described in Section 1, and that the quantity corresponding to Y may be $(Z-X)$ dB, the difference between X and Z at any time.

Instead of the correlation between X and Y, let us find the correlation between X and $Z-X$. Figs. 21 and 22 are few examples of the correlation. Then little correlation can be found between them over a long-period as shown in Figs. 21 and 22, even if positive or negative correlation is found between them for a short-period. In the consideration made over a fairly long-period, x and y may be thought of as approximately independent of each other.

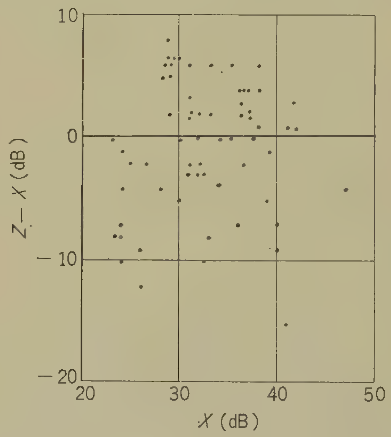


Fig. 21—Scatter diagram between $Z-X$ and X.

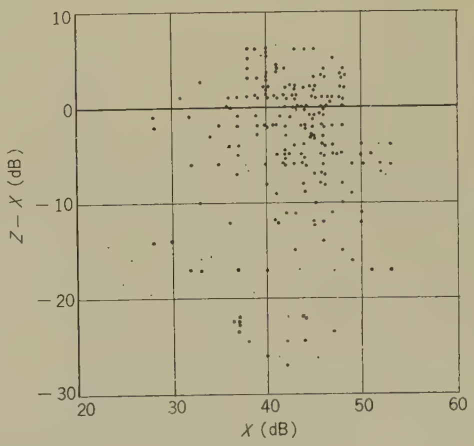


Fig. 22—Scatter diagram between $Z-X$ and X.

If the instantaneous value of long-period variation, x, and that of short-period variation, y, are assumed to follow the following assumption:

(a) The values of x and y in dB have Gaussian distribution.

(b) x has Gamma-distribution, and y Beta-distribution.

(c) x and y have Gamma-distribution.

(d) The value of x in dB has Gaussian distribution, and y, Gamma-distribution,

then the distribution function of z (random product of x and y) becomes as follows by calculation:

(a) The distribution of z in dB for condition (a) is Gaussian distribution.

(b) The distribution of z for condition (b) is Gamma-distribution.*

(c) The distribution of z for condition (c) becomes the sums of Gamma-distribution.**

(d) The distribution which z follows for condition (d) becomes after effect function.⁷⁾***

According to actual measurements, the distribution of the instantaneous value z of the received power measured over various propagation path for periods of 10 days to one month can be approximated by one of following distributions:

(1) The dB instantaneous value of the received power intensity follows Gaussian distribution.

(2) The instantaneous value of the received power intensity follows Gamma-distribution.

(3) The instantaneous value of the received power intensity follows the composite distribution of (1) and (2).

As a result, the assumption of the distribution function of x and y is considered not too unreasonable, and so the composition of fading is decided to be the product of x and y.

3.2. Frequency Characteristics

When the long-period variation level X and the short-period variation level Y follow respectively the Gaussian distribution $N\ (M_X, \sigma_X)$ and $N\ (0, \sigma_Y)$ the product of x and y (the product being shown by z) follows the Gaussian distribution $N\ (M_Z, \sigma_z)$. That is, the following relations are obtained:

$$\sigma_Z = \sqrt{\sigma_X{}^2 + \sigma_Y{}^2}, \qquad M_Z = M_X$$

If the variation ranges of X, Y, and Z against the 1–99% probability, i.e. fading ranges, are shown by F_0, F_f, and F_{rf} respectively, the relation:

$$F_{rf} = q\sqrt{\sigma_X{}^2 + \sigma_Y{}^2} = \sqrt{F_0^2 + F_f^2}$$

is obtained.
Where

$$q\sigma_X = F_0, \quad q\sigma_Y = F_f, \quad q = 4.66.$$

The fading ranges F_{rf} evaluated by the exact formula for conditions (b), (c) and (d) in Section 3.1 are plotted in Figs. 23, 24, and 25 against the fading ranges evaluated by

$$f_0(x) = \frac{a^{\nu+1}}{\Gamma(\nu+1)} x^{\nu} e^{-ax}, \quad f_f(y) = \frac{1}{B(n_1, n_2)} y^{n_1-1}(1-y)^{n_2-1}$$

$$f_{rf}(z) = \int_0^{\infty} f_0(x) f_f(y) \frac{\partial(x, y)}{\partial(x, z)}\, dx \qquad \text{where} \quad z = x.\,y$$

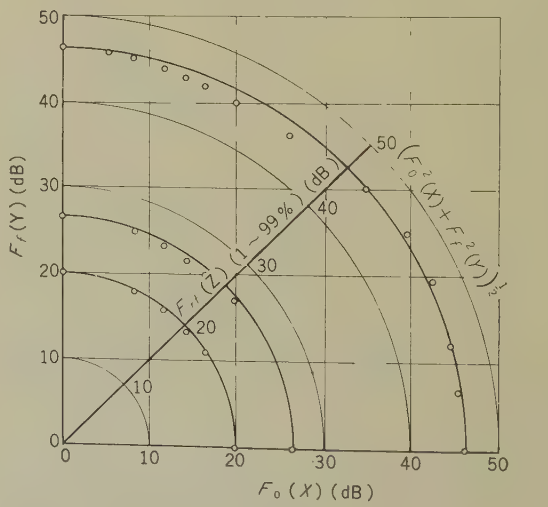


Fig. 23—Deviation of $\{F_0^2(X)+F_f^2(Y)\}^{1/2}$ from exactly evaluated value $F_{rf}(Z)$.

---

* See appendix (i). ** See appendix (ii).
*** See appendix (iii).

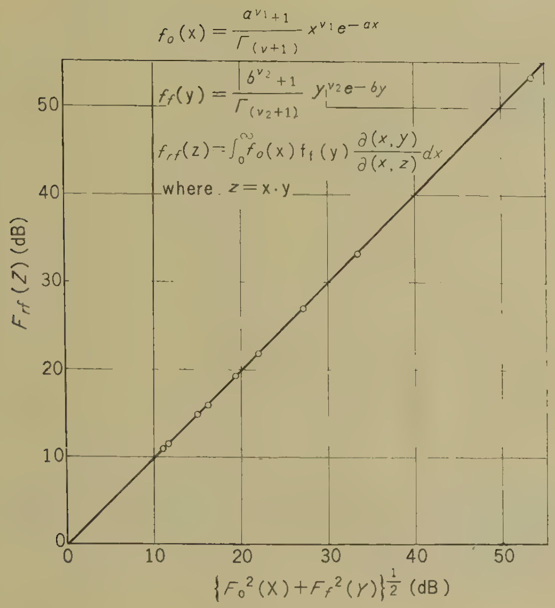


Fig. 24—Deviation of $\{F_0^2(X)+F_f^2(Y)\}^{1/2}$ from exactly evaluated value $F_{rf}(Z)$.

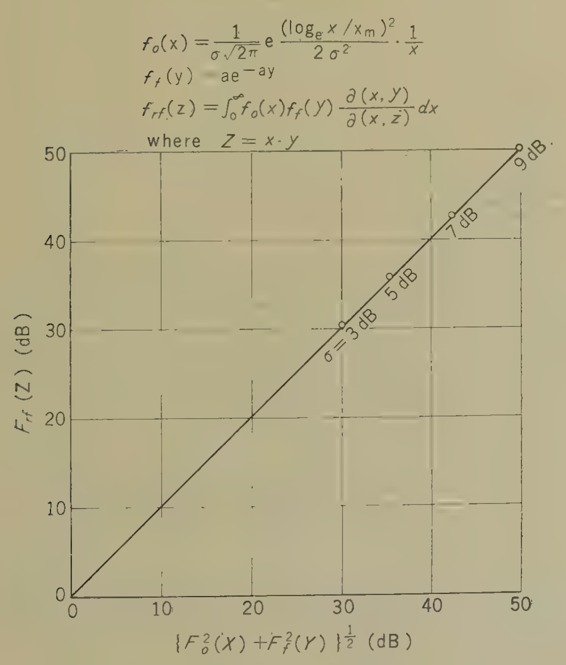


Fig. 25—Deviation of $\{F_0^2(X)+F_f^2(Y)\}^{1/2}$ from exactly evaluated value $F_{rf}(Z)$.

the following approximation formula $F_{rf}=\sqrt{F_0^2+F_f^2}$.

The errors due to the values evaluated by $F_{rf}=\sqrt{F_0^2+F_f^2}$ are less than 2 dB.

Therefore the fading range can be approximately represented by the following equation:

$$F_{rf}=\sqrt{F_0^2+F_f^2} \tag{16}$$

Next, F_f is written in the form,

$$F_f=(\beta f)^p \tag{17}$$

where p and β are constant coefficients being independent of frequency f. Then the following relation is obtainable:

$$F_{rf}\equiv\sqrt{F_0^2+(\beta f)^{2p}} \tag{18}$$

If the fading ranges are measured for more than three given frequencies in the same path and the same period, three unknown factors p, β and F_0 in Eq. (18) can be evaluated.

Table 4 shows the values of p obtained from the simultaneous experiments on the same propagation paths. The frequency characteristics of fading range of short-period variation are found to be approximately proportional to $f^{1/2}$. Let the value of p to be $p=0.5$, and the following expression is possible:

$$F_{rf}=\sqrt{F_0^2+\beta f} \tag{19}$$

where the unit of f is Gc.

If the values of fading range, F_{rf} are measured simultaneously for two frequencies on the same path, F_0 and β of the path are obtained by Eq. (19). The correlation between F_0 and β obtained in this way is shown in Fig. 26. The correlation coefficient, $r=0.972$ is obtained.

According to the F-test, the correlation can be regarded significant for sample $N=8$. If the regression line of β to F_0 is assumed to be

$$\log_{10}F_0=m\log_{10}\beta+l,$$

Table 4

OBSERVED VALUE OF p

	Month and year	Frequency (Gc)	p	$\bar{p}$
Mt. Futago–Mt. Tsukuba	June, 1954	1.5, 4, 6	0.43	0.508
〃	July～Aug. 1954	0.2, 1.5, 6	0.52	
〃	July～Aug. 1954	1.5, 4, 6	0.49	
Fudo-tôge–Itsukaichi	Aug.～Sep. 1956	1.5, 4, 6.7	0.55	
Kawaguchi–Itsukaichi	July～Aug. 1956	1.5, 4, 6.7	0.55	

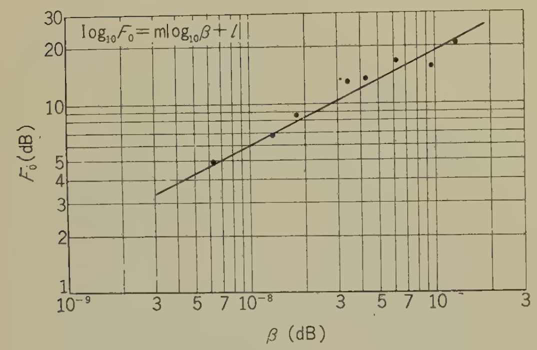
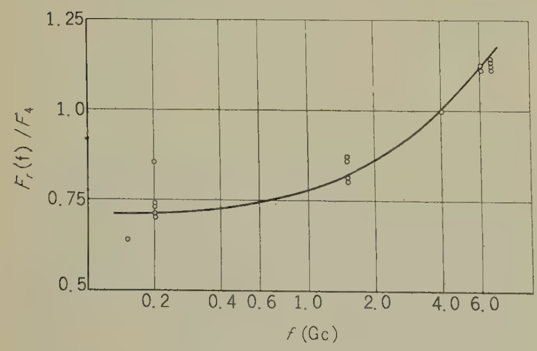


Fig. 26—Correlation between F_0 and β.



Solid Line : Theoretical Value

Fig. 27—Frequency characteristics of fading range.

the regression coefficient m becomes 0.464. For convenience, the next approximation is adopted:

$$\beta=aF_0^2$$

If "a" is evaluated by the method of least squares, it becomes:

$$a=0.28\pm0.03 \qquad (20)$$

Eq. (18) is expressed, more generally, as follows:

$$\delta_{i,j}=F_{r,i,j}-\sqrt{F_{0,i}+Kf_{i,j}F_{0,i}^{2p}} \qquad (21)$$

where i=propagation path, j=frequency.

The value of "a" in Eq. (20) nearly corresponds to that of "a" evaluated by the method of least squares as for Eq. (21).

The error due to using $\beta=0.28F_0^2$ is within about 5% over the variation range of F_0. As a result, the following form will be adopted as the frequency function of fading range.

$$F_{rf}=F_0\sqrt{1+0.28f} \qquad (22)$$

where F_0=The characteristic constant of the propagation path for the season.
f=Frequency in Gc/s.

The normalized frequency characteristics of fading range at 4 Gc are shown in Fig. 27

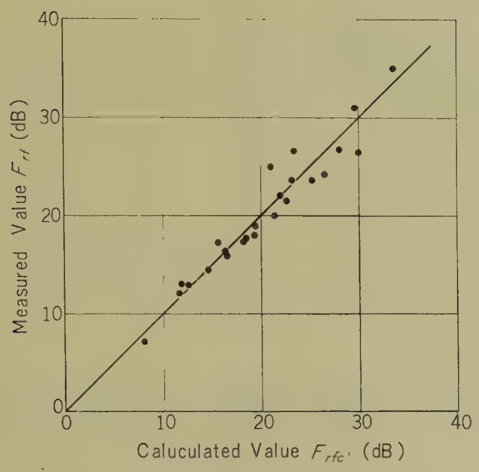


Fig. 28—Comparison between measured fading range and calculated one.

The solid line in the figure corresponds to Eq. (22). Fig. 28 shows the comparison between the measured fading ranges of various frequencies at the same time and at the same path, and the fading ranges evalated by Eq. (22).

F_{rf} can also be written as

$$\left.\begin{aligned} &F_{rf}=C_fF_0=\sqrt{F_0^2+afF_0^2}=\sqrt{F_0^2+F_f^2}\\ &\text{Accordingly, } F_f=\sqrt{a}\,f^{1/2}F_0. \end{aligned}\right\}\quad(23)$$

Therefore, F_{rf} and F_f are proportional to F_0.

This relation has been obtained by qualitative estimation from the experimental characteristics of Section 1.4. One of the experimental results which show its propriety will be described in the following.

Figs. 29 and 30 show the records of 4 Gc propagation tests made over two paths—path A and path B—(See Fig. 35) between the middle side of Tsukuba Hill and the broadcasting tower at Kawaguchi with receiving antennas of different heights. These are examples which show diurnal variation ranges in dB (max., min./5 min.).

In these figures, (a) shows an example for a day with heavy fading, while (b) an example for a day with light fading. At a glance it is found that when fading range F_{rf} is large the long-period variation varies more and its duration is longer than when F_{rf} is small. It is supposed that the proportional relation between F_{rf} and F_0 shown by Eq. (23) is little related to the receiving height.

As for these two paths, the daily fading range of F_0 was derived from the daily fading range of F_{rf} by using the method shown in Section 1.3 and the correlation between these daily fading ranges is plotted in Fig. 31.

The correlation coefficient between these ranges is larger than 0.84 (this value being

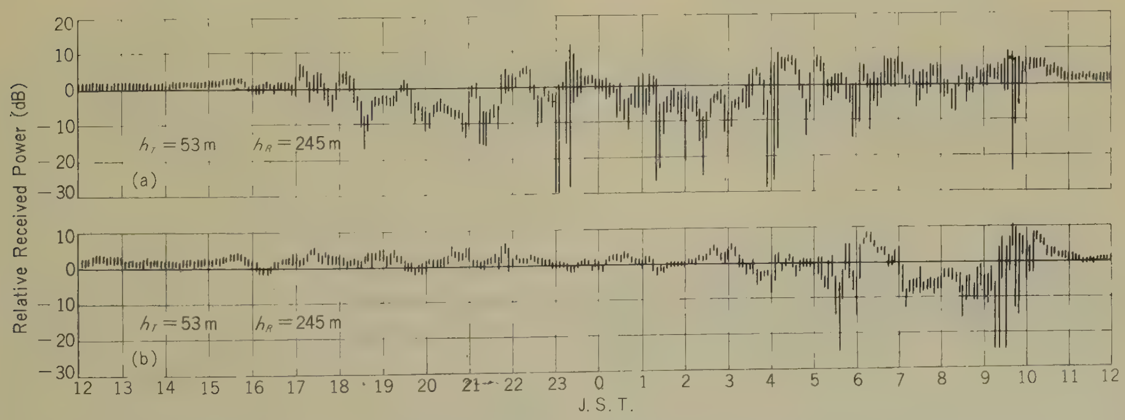


Fig. 29—Examples of diurnal signal variation received on the Kawaguchi–Mt. Tsukuba path at 4 Gc/s. (a) A day (Nov. 9~10, 1954) with heavy fading; (b) A day (Nov. 6~7, 1954) with light fading.

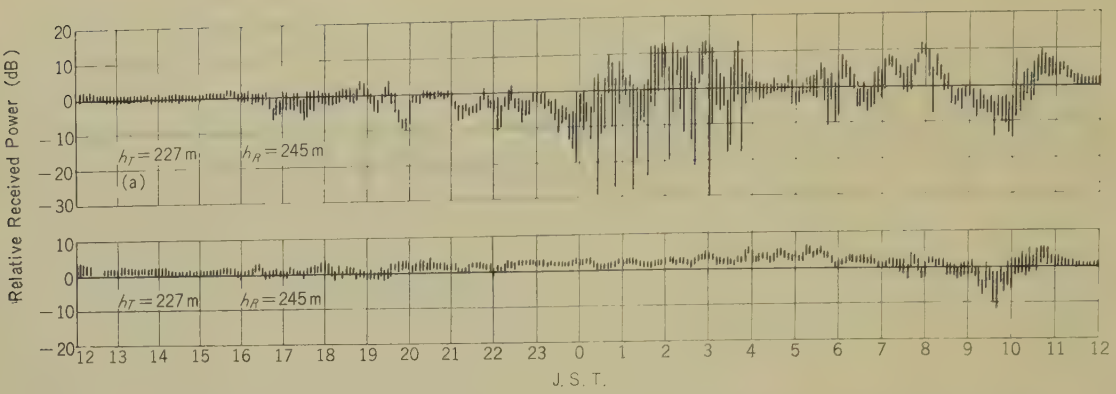


Fig. 30—Examples of diurnal signal variation received on the Kawaguchi–Mt. Tsukuba path at 4Gc/s. (a) A day (Nov. 9∼10, 1954) with heavy fading; (b) A day (Nov. 6∼7, 1954) with light fading.

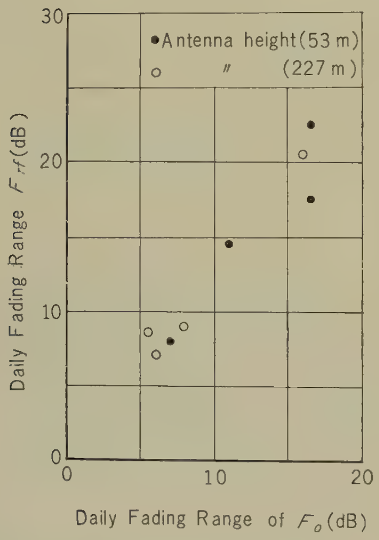


Fig. 31—Correlation between daily range of F_{rf} and daily fading range F_0 obtained on both path A and B. Path A and B are indicated by mark ● and ○ respectively.

the correlation coefficient at the 1% significance level by the Chi-test).

As a result, it is considered that F_{rf} is proportional to F_0. In the figure, the mark ● shows the value for a path with a low receiving point. This proportional relation shows that it is hardly influenced by the difference of receiving height.

3.3. Seasonal Characteristics of Fading

In Section 1.4.1, the outline of the correspondences between the fading ranges and various meteorological factors are shown.

Among these correspondences, it is found that fading ranges are closely related to the parameter which includes the variance of average gradient of modified refractive index M of the air. The reason is thought to be as follows:

According to the consideration in Section 2.2.2, the variation of long-period level is caused by the convergence and divergence of radio energy due to the bending part of the M profile, and the fading range F_0 of long-period variation is written as

$$F_0 \propto \sigma_X$$

Therefore from Eq. (13)

$$F_0=\mathrm{func\,n}\left\{\frac{\sigma_{\alpha_0}\cdot\sqrt{\Delta h_0}}{\langle\alpha\rangle}\right\} \qquad (24)$$

where σ_{α_0} shows the standard deviation of average gradient α_0 of the modified refractive index of the stratified layer with the thickness Δh_0.

The following experimental result is given in reference.[6)]

$$\sigma_{\alpha_0}^2\cdot\Delta h_0=\sigma_\alpha^2\cdot\Delta H=\text{invariant}$$

where σ_α shows the standard deviation of average gradient α of the modified refractive index of the layer with the thickness ΔH.

The sign $\langle\ \ \rangle$ means timely mean value.

Therefore if ΔH is assumed to be a standard value, the following relation can be written:

$$F_0=\text{func}\,\underline{n}\left\{\frac{\sigma_\alpha}{\langle\alpha\rangle}\right\} \quad \text{or} \quad F_{rf}=C_f\cdot\text{func}\,\underline{n}\left\{\frac{\sigma_\alpha}{\langle\alpha\rangle}\right\} \tag{25}$$

C.C.I.R's advice proposes that $\Delta H=1000$ m (1000 m above the ground) should be taken. The value of M on the ground, however, is easily controlled by local meteorological phenomena, influenced by objects on the ground, and influenced by topography.

Accordingly, it is doubtful whether it is suitable or not to insert this into the amount which estimates the state of the upper layer. Besides, practically it is not easy to evaluate accurately the meteorological condition at the appointed altitude of 1000 m from the surface. From these points of view, radiosonde data, which are found in the Aerological Data published monthly from the Meteorological Agency are used. And the atmospheric pressure level of 1000 mb (about 100 m above the ground) is used instead of the ground surface. Moreover, instead of the altitude of 1000 m, the appointed atmospheric pressure level 900 mb (about 1000 m above sea level) is adopted. The mean value $\langle\alpha\rangle$ of the average gradient α between the two pressure levels and the standard deviation σ_α for a month are calculated. These calculated values are regarded as the representative values for the month.

If the form of the function of Eq. (25) is found, the seasonal movement of the fading range can be brought to light.

3.3.1. Experimental Equation of the Function func $\underline{n}\left\{\frac{\sigma_\alpha}{\langle\alpha\rangle}\right\}$

The fading range data are consisted of the results measured in the path between Fudōtōge and Itsukaichi, which runs parallel to the ground surface. In the scatter diagram a in Fig. 32, the correlation of $\sigma_\alpha/\langle\alpha\rangle$ obtained from the radiosonde data at Tateno during the same period in the month is shown against the monthly fading ranges at the frequency of 4 Gc, which were measured throughout one year over the same path. In this figure, u is used instead of σ_α. Here, considering the small number of the samples of α, unbiased variance u_α^2 is used. As shown in the figure, a very good linear correlation is obtained.

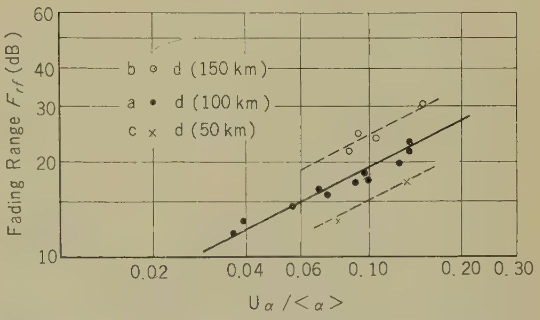


Fig. 32—Correlation between fading range F_{rf} and $u_\alpha/\langle\alpha\rangle$ with respect to three paths.

So, the function form of Eq. (25) is as follows:

$$F_{rf}=C\cdot\left\{\frac{u_\alpha}{\langle\alpha\rangle}\right\}^q \tag{26}$$

Then the constants q and c are obtained by the method of least squares,

$$\left.\begin{aligned} q&=0.4445 \\ C&=53.93 \end{aligned}\right\} \tag{27}$$

The calculated correlation coefficient ρ between the measured fading range F_r (4 Gc) and the fading range F_r' (4 Gc) estimated from Eqs. (26) becomes:

$$\rho=0.97.$$

The square value s^2 of the difference between measured value and the estimated value is obtained from

$$s^2=0.645\ (\text{dB})^2$$

Accordingly,

$$s=\pm 0.804\ \text{dB}$$

For the convenience of practical calculation, if the method of least squares is used with the assumption $q=0.5$, the following is obtained:

$$C=60.4$$

The seasonal variation range of $\sigma_\alpha/\langle\alpha\rangle$ through a year is usually less than the range from 0.01 to 0.25, as shown in Fig. 32 and Table 6. The error due to the replacement of q from 0.4445 to 0.5 is found to be within a range of ± 5 % at the most as the result of calculation. Accordingly, the index q is chosen to be 0.5 for the convenience of practical calculation. The solid line a in Fig. 32 shows the estimated value when q is equal to 0.5.

3.4. Distance Characteristics of Fading

By combining Eq. (22) and the result above, the fading range F_{rf} for any frequency can be expressed as:

$$F_{rf}\propto\sqrt{1+0.28f}\left\{\frac{\sigma_\alpha}{\langle\alpha\rangle}\right\}^{1/2}\cdot\text{func }\underline{\text{n}}(d)\cdot\phi(h_t,h_r) \tag{28}$$

Where func $\underline{\text{n}}$ (d) and $\phi(h_t, h_r)$ respectively show the functions with the variables of distance and height of the transmitting (receiving) antenna above the ground. $\phi(h_t, h_r)$ will be dwelt upon later. When $\langle\alpha\rangle$ and σ_α are assumed constant with altitude, it is considered from Section 2.2.4 that fading range becomes largest in case of $h_t=h_r$ and $\phi(h_t, h_r)=1$. The analysed paths are the path between Fudō-tōge and Itsukaichi (97 Km), the path between Mt. Tsukuba and Mt. Futago (147 Km), and the path between Kawaguchi and Itsukaichi (48 Km). The characteristics of these paths (relative position of the path to the ground) are not the same, but the assumption $\phi\fallingdotseq 1$ is thought not to cause a large error, as a result of consideration of the altitude characteristics of $\langle\alpha\rangle$ and σ_α, and consideration of the propagation path height, and of its inclination against the mean surface of the earth which will be described in Section 3.6.

In Fig. 32, the data for each of the above path are shown by a, b, and c respectively.

There was no experiment made over a complete year. The monthly fading ranges during the measuring period are plotted against the values of $U_\alpha/\langle\alpha\rangle$ obtained from the radiosonde at Tateno during the period.

The measuring points of both b and c run nearly parallel to the measuring points of a.

This fact shows that the law of $\{\sigma_\alpha/\langle\alpha\rangle\}^{1/2}$ can be applied for every distance.

3.5. Comparison between Experimental Formula of Fading Range Estimation and Its Measured Value

After eliminating the frequency effect and seasonal effect of fading range, the distance characteristics of F_0 (d) are as shown in Fig. 33.

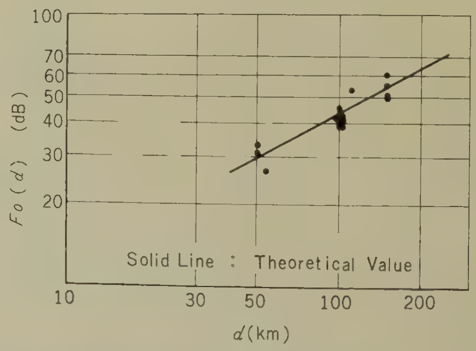


Fig. 33—Characteristics of F_0 against path distance.

Where

$$F_0(d)\equiv\text{func }\underline{\text{n}}(d)=\frac{F_{rf}}{\sqrt{1+0.28f}\left\{\frac{\sigma_\alpha}{\langle\alpha\rangle}\right\}^{1/2}}\cdot \tag{29}$$

A fairly good linear correlation is obtainable. Therefore the func n (d) is shown by d^n, and the fading range is written as

$$F_{rf}=K\sqrt{1+0.28f}\left\{\frac{\sigma_\alpha}{\langle\alpha\rangle}\right\}^{1/2} d^n \text{ (dB)}. \qquad (30)$$

Finally, if K and n are evaluated by the method of least squares, the following experimental formula is obtained:

$$F_{rf}=3.73\sqrt{1+0.28f}\left\{\frac{\sigma_\alpha}{\langle\alpha\rangle}\right\}^{1/2} d^{0.53}\pm\varepsilon \text{ (dB)} \qquad (31)$$

$$\varepsilon=1.61 \text{ (dB)}.$$

Where ε is the root mean square of the difference between the measured value and calculated value.

In Fig. 34, a comparison between the calculated and measured values is shown. From this, fading range can be considered proportional to the 0.53 power of the distance.

3.6. The Effects of Path Height and its Inclination

3.6.1. The Influence of Path Height upon Fading Range

The value of $\phi(h_t, h_r)$ is considered to be determined mainly by the average height of propagation path above the ground and the inclination of the path itself to the ground.

When the path inclination is fixed, the influence of the average height upon fadings is dependent on the characteristics of the gradient of the refractive index of the air and its variance in every season and in every region.

For instance, the result of the variance analysis in every district as to σ_α obtained from radiosonde data in 1951 is shown in Table 5.

From this table, σ_α is found to be related to height h (0–500 m), (500–1000 m), (1000–1500 m) and month m (seasonal effect), but the difference of σ_α against the factor h at all locations is not significant on the 1% significance level.

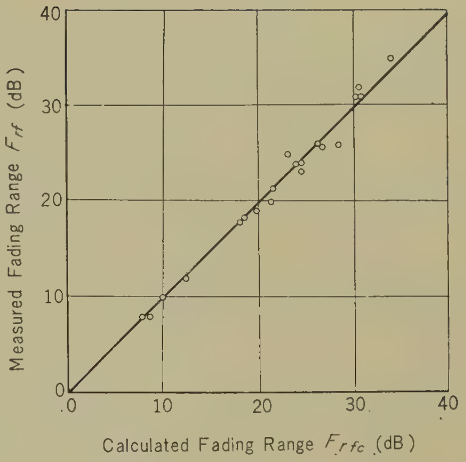


Fig. 34—Comparison between measured fading range and calculated one by estimation formula.

According to the result of statistical analysis of M profiles in the layers more than 30 m above the ground (in layers less than 30m above the ground, the data are not sufficient), the characteristics of the mean value of the modified refractive index gradient and the characteristics of its standard deviation against height are not so remarkable below 1500 m above the ground at the nine radiosonde observatories in Japan.[6)]

If the above stated characteristics can be regarded as general, there seems to be not such a large influence, if any, of the mean height of propagation path upon the fading range. On the basis of the significance test, the effects of the inclination of the path upon fading are considered larger.

3.6.2. The Influence of Path Inclination upon Fading Range, and the Characteristics of ϕ (h_t, h_r)

In general, the surface of the ground is not plain. Besides, the propagation path is not usually linear. By using the mean gradient $\langle\alpha\rangle$ (M.U./m) of the M profile defined within the range of the radio propagation path, the radius R of the curvature of the

Table 5

RESULT OF ANALYSIS OF VARIANCE

District	Result of analysis of variance				
	Factor	Variation	Freedom	Unbiased variance	Judgment
Wakkanai	m	473. 3	11	43. 1	*
	h	21. 6	2	10. 8	
	$m \times h$	63. 1	22	2. 87	
Sapporo	m	344. 9	11	31. 4	*
	h	24. 2	2	12. 1	
	$m \times h$	65. 8	22	2. 99	
Akita	m	399. 9	11	36. 4	*
	h	13. 5	2	6. 75	
	$m \times h$	29. 8	22	1. 35	
Sendai	m	438. 3	11	39. 8	*
	h	6. 1	2	3. 05	
	$m \times h$	30. 6	22	1. 39	
Haneda	m	360. 2	11	32. 7	*
	h	7. 7	2	3. 85	
	$m \times h$	59. 0	22	2. 68	
Wajima	m	565. 6	11	51. 4	*
	h	9. 0	2	4. 50	
	$m \times h$	99. 7	22	4. 53	
Shionomisaki	m	543. 0	11	49. 4	*
	h	34. 9	2	17. 45	
	$m \times h$	75. 1	22	3. 41	
Yonago	m	427. 6	11	38. 9	*
	h	12. 0	2	6. 0	
	$m \times h$	40. 7	22	1. 85	
Kagoshima	m	183. 0	11	16. 9	*
	h	6. 0	2	3. 0	
	$m \times h$	61. 3	22	2. 79	

Notes: Numerical value is multipled by 500.
* Significant. (1%)

radio path is given as follows:

$$R=\frac{10^6}{\left\langle \dfrac{dM}{dh} \right\rangle}=\frac{10^6}{\langle\alpha\rangle} \quad (\mathrm{m})$$

Even if transmitting and receiving antenna heights are not equal, a part of the propagation path will be parallel to the ground surface. The condition for this is shown by

$$d \geq \sqrt{\frac{2(h_t \sim h_r)}{\langle\alpha\rangle}}=1.41\sqrt{\frac{h_t \sim h_r}{\langle\alpha\rangle}} \quad (\mathrm{km}) \quad (32)$$

where $h_t(h_r)$: The height of transmitting (receiving) antenna

In this case, from the result of geometrical analysis in Section 2.2.4, it is considered that the contribution of path characteristics to fading range is most remarkable in the part of the propagation path which is in parallel to the ground surface. The fading range taken in this part seems to determine the greater part of the fading range of the path. Therefore, under the condition of Eq. (32), the quantitative characteristics of the fading range is determined by the fading range taken in this part, even when h_t is not equal to h_r.

In the following, an experiment to show that the above is appropriate, will be described.

The two propagation experiments were carried for the distance of 54.8 Km between Tsukuba Hill and Kawaguchi. The transmitting antenna height at Tsukuba Hill was 245 m above sea level and the receiving antennas at Kawaguchi were installed at two different heights, 53 m and 227 m above sea level. We call the former path and the latter path A and B respectively. On both of these propagation path, the reflected waves on the ground are almost shielded by standing trees or topography near the reflection points as shown in Fig. 35.

Accordingly, large fading due to the interference of the ground reflection wave can be eliminated, and then the fading due to residual ground reflection waves can be ignor-

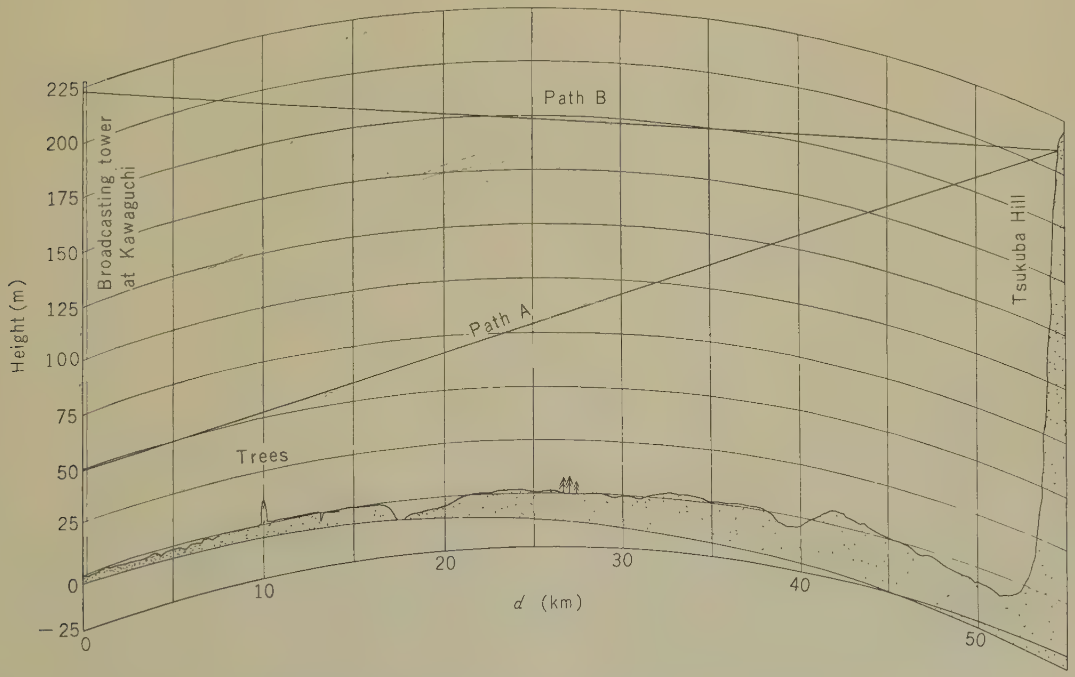


Fig. 35—Path profile from Kawaguchi to Tsukuba-Hill.

ed as compared with the fading range due to ducts, because the fading range is defined as the variation width of 1%–99%.

Radio frequencies used for tests in both paths are different from each other by only 0.04 Gc/s in the 4 Gc/s band, and so the difference of fading range caused by the difference of frequency is not significant in the discussion. Therefore, the experiment can be said to be a comparison of the profiles of these paths. The ground along these paths is nearly flat and rises gradually in the direction of Tsukuba Hill.

The structure of the atmosphere is thought to be stratified in parallel to the ground surface, and so the profile of the propagation path in the figure is drawn by taking the mean ground surface as the standard of the ground level. The height of the transmitting antenna is about 210 m above this standard level, and the condition of Eq. (32) is satisfied concerning both the paths.

Fig. 36 shows the quantitative comparison of the diurnal fading ranges. On days with small fading range in path B, the fading range in the so-called inclined path A becomes larger. This is considered to be due to the irregularity of the M profile which is severer near the ground surface than in the upper layer.

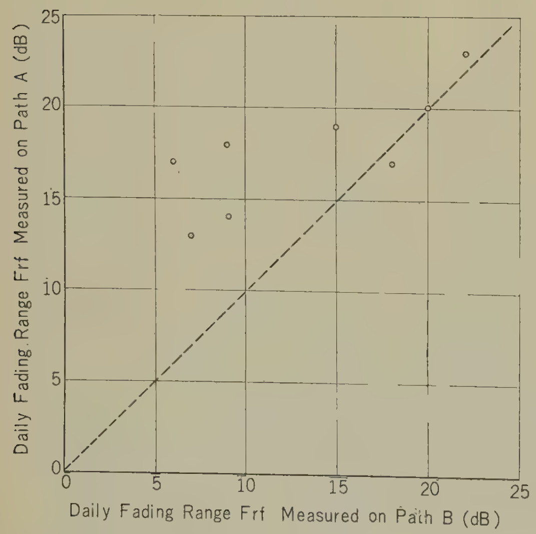


Fig. 36—Comparison of daily fading range measured on path A with that measured on path B.

But on days with large fading range, the ranges in both the paths are nearly equal, with little difference caused by the inclination of the path.

Such characteristics are obtained, probably, because a part of the path which looks inclined is parallel to the ground surface.

3.6.3. Path Inclination and the Inclination Coefficient

$$d<1.41\sqrt{\frac{h_t\sim h_r}{\langle\alpha\rangle}}\quad(\text{km})\qquad(33)$$

In order to satisfy the condition of Eq. (33) the radio path does not run parallel to the ground. The path which satisfies this condition, is defined as a slanting path in this paper, while that which satisfies Eq. (32) is regarded as a horizontal path. Moreover, in this case, $\phi(h_t, h_r)$ is assumed to be 1.

As supposed from Section 2.2.4, fading range varies with the inclination of the radio path to the ground.

Let $\phi(h_t, h_r)$ be the ratio of the measured fading range in a slanting path to the fading range obtained from Eq. (31).

We shall introduce the parameter u which shows the inclination of path to the ground, and put it as follows:

$$\left.\begin{aligned} u&=\frac{\langle\alpha\rangle d}{2\theta},\\ &\text{where}\\ \theta&=\frac{h_t\sim h_r}{d}. \end{aligned}\right\}\qquad(34)$$

If u is plotted against $\phi(h_t, h_r)$ by using Eq. (34) for experiments, the relation in Fig. 37 is obtained. Here, from the approximate tendency between $\phi(h_t, h_r)$ and u, $\Phi(u)$ can be found.

The fading range in a slanting path can be represented by $\Phi(u)$. Therefore $\Phi(u)$ is

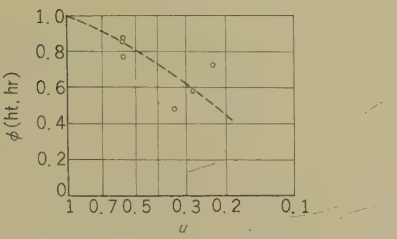


Fig. 37—Relation between inclination coefficient $\Phi(u)$ and parameter u.

called the inclination coefficient of a path.

3.7. Estimation Formula for Fading Range and Its Application

3.7.1. The Outline of the Regional Distribution of the Refractive Index of the Air in Japan

Japan faces the Siberian Continent over the Sea of Japan to the north, the Pacific Ocean to the south, and most of the land is occupied by the high mountains running lengthwise through the middle.

Therefore the meteorological conditions are very complicated. Accordingly, the atmospheric conditions relating to radio propagation have regional differences quite remarkable.

Especially remarkable is the effect of latitude difference. Typical examples of extreme regions in Japan are shown in Fig. 38.

These examples were calculated monthly from the aerological data at Kagoshima, Tateno and Sapporo during the 3 years from 1954 to 1956.

The values of $\langle\alpha\rangle$ is monthly mean of the average gradient of modified refractive index and is calculated per unit height between the two altitudes whose air pressure are 1000 mb and 900 mb. The value σ_α is the standard deviation of α in the same period. In a later section, these factors will be called atmos-

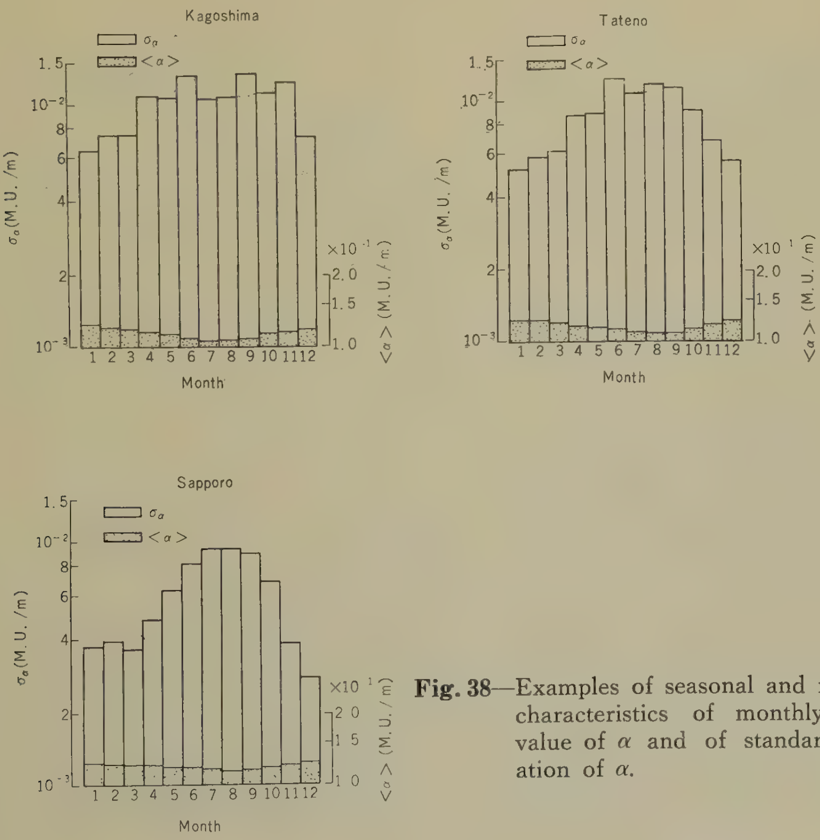


Fig. 38—Examples of seasonal and regional characteristics of monthly mean value of α and of standard deviation of α.

Table

VALUES OF $\langle\alpha\rangle$ AND σ_α BETWEEN 1000 MB

	Wakkanai		Sapporo		Akita		Sendai	
	$\langle\alpha\rangle$	σ_α	$\langle\alpha\rangle$	σ_α	$\langle\alpha\rangle$	σ_α	$\langle\alpha\rangle$	σ_α
JAN.	0.1227	0.00276	0.1214	0.00370	0.1206	0.00330	0.1209	0.00336
FEB.	0.1220	0.00441	0.1210	0.00395	0.1208	0.00342	0.1223	0.00522
MAR.	0.1210	0.00324	0.1212	0.00363	0.1206	0.00439	0.1203	0.00602
APR.	0.1183	0.00531	0.1202	0.00483	0.1178	0.00683	0.1178	0.00688
MAY	0.1145	0.00832	0.1170	0.00653	0.1131	0.00821	0.1146	0.00916
JUN.	0.1144	0.00808	0.1161	0.00818	0.1112	0.00774	0.1133	0.00996
JUL.	0.1146	0.01151	0.1147	0.00943	0.1098	0.00907	0.1090	0.01092
AUG.	0.1128	0.01329	0.1112	0.00945	0.1077	0.01111	0.1084	0.01161
SEP.	0.1101	0.01185	0.1128	0.00798	0.1111	0.01100	0.1111	0.01167
OCT.	0.1170	0.00713	0.1173	0.00675	0.1169	0.00698	0.1170	0.00817
NOV.	0.1213	0.00429	0.1206	0.00381	0.1193	0.00574	0.1195	0.00534
DEC.	0.1277	0.00522	0.1218	0.00277	0.1202	0.00397	0.1206	0.00356

pheric radio meteorological constants.

The value of $\langle\alpha\rangle$ at Sapporo is larger than that at Kagoshima all year round, and $\langle\alpha\rangle$ at Tateno is smaller than that at Sapporo but larger than that at Kagoshima. Its seasonal value, however, is least in summer, largest in winter, and its deviation is small at every place.

On the other hand, the regional difference of σ_α is larger than that of $\langle\alpha\rangle$, that is, the value of σ_α is much larger at Kagoshima than at Sapporo on the whole. The value of σ_α at Tateno is midway between them.

The seasonal variation of σ_α is also contrary to the variation of $\langle\alpha\rangle$. That is, it is largest in summer, and smallest in winter. Besides, the variation is far greater than that of $\langle\alpha\rangle$, and that in summer is over twice as large as that in winter.

Table 6 shows the monthly atmospheric radio meteorological constants between 1954 to 1956 calculated from radiosonde data at nine Aerological Observatories in Japan.

Table 7 shows an example of comparison between the constants over the sea and those over land.

Table 7

COMPARISON OF VALUES OF $\langle\alpha\rangle$ AND σ_α BETWEEN ON LAND AND OVER SEA

1957	Torishima		Tateno	
	$\langle\alpha\rangle$	σ_α	$\langle\alpha\rangle$	σ_α
Aug.	0.094	0.0091	0.100	0.0087
Sep.	0.101	0.0168	0.109	0.0116

Between fading ranges and the atmospheric radio meteorological constants obtained by radiosonde observation, as already mentioned, it has been proved that there is a good correlation from experimental and theoretical considerations. Therefore by using data from radiosonde observations, the fading range everywhere in Japan can be estimated.

6

AND 900 MB LEVELS (1954~1956)

Tateno		Wajima		Shionomisaki		Yonago		Kagoshima	
$\langle\alpha\rangle$	σ_α	$\langle\alpha\rangle$	σ_α	$\langle\alpha\rangle$	σ_α	$\langle\alpha\rangle$	σ_α	$\langle\alpha\rangle$	σ_α
0. 1206	0. 00522	0. 1197	0. 00434	0. 1220	0. 00483	0. 1194	0. 00394	0. 1203	0. 00647
0. 1209	0. 00593	0. 1196	0. 00430	0. 1206	0. 00516	0. 1194	0. 00349	0. 1190	0. 00736
0. 1193	0. 00617	0. 1186	0. 00538	0. 1172	0. 00769	0. 1182	0. 00552	0. 1182	0. 00747
0. 1157	0. 00872	0. 1161	0. 00829	0. 1119	0. 11116	0. 1158	0. 00906	0. 1136	0. 01095
0. 1125	0. 00891	0. 1119	0. 00973	0. 1070	0. 01143	0. 1115	0. 00971	0. 1110	0. 01060
0. 1114	0. 01243	0. 1080	0. 00970	0. 1025	0. 01183	0. 1087	0. 01129	0. 1071	0. 01307
0. 1065	0. 01071	0. 1058	0. 01184	0. 0964	0. 01341	0. 1055	0. 01190	0. 1051	0. 01046
0. 1075	0. 01195	0. 1041	0. 01288	0. 0974	0. 01072	0. 1021	0. 01046	0. 1062	0. 01077
0. 1091	0. 01137	0. 1070	0. 01218	0. 1041	0. 01271	0. 1079	0. 00982	0. 1088	0. 01346
0. 1145	0. 00928	0. 1130	0. 00915	0. 1134	0. 00846	0. 1141	0. 00684	0. 1136	0. 01117
0. 1174	0. 00690	0. 1176	0. 00831	0. 1167	0. 00806	0. 1181	0. 00469	0. 1146	0. 01230
0. 1212	0. 00572	0. 1190	0. 00420	0. 4196	0. 00663	0. 1195	0. 00401	0. 1181	0. 00733

3.7.2. Method of Estimation of Fading Range and Its Application

If the path distance d in km, the transmitting (receiving) antenna height above mean ground level h_t in m (h_r in m), and the radio frequency f in Gc are given; the fading range F_{rf} (dB) can be obtained as follows by finding the atmospheric radio meteorological constanst $\langle\alpha\rangle$ and σ_α:

$$\left.\begin{aligned} &F_{rf}=3.73\sqrt{1+0.28f}\left\{\frac{\sigma_\alpha}{\langle\alpha\rangle}\right\}^{1/2} d^{0.53}\Phi(u),\\ &\text{where}\quad u=\frac{\langle\alpha\rangle d}{2\theta}=\frac{\langle\alpha\rangle d^2}{2(h_t\sim h_r)} \end{aligned}\right\} \tag{35}$$

$\sigma_\alpha/\langle\alpha\rangle$: Radio meteorological coefficient determined by σ_α and $\langle\alpha\rangle$.

$\Phi(u)$: Path inclination coefficient determined by u.

For the convenience of calculation, nomographs for Eq. (35) are shown on the assumption $\Phi(u)=1$ in Fig. 39. The approximate value of $\Phi(u)$ is found from Fig. 37.

Table 8 shows a comparison of fading ranges measured recently (August, 1958) for a microwave circuit between Toyama and Asahi with those calculated by Eq. (35) from the atmospheric radio meteorological coefficients obtained from radiosonde data for the same period at Wajima. Fairly good correspondence is found.

Table 8

COMPARISON BETWEEN THE MEASURED VALUES AND THE CALCULATED VALUES OF FEDING RANGE

f (Gc) / F_{rf} (dB)	1. 5	4	6. 7	Φ	d (km)
Calculated Value	9. 7	11. 8	13. 7	0. 84	44
Measured Value	10	11. 5	12. 5		

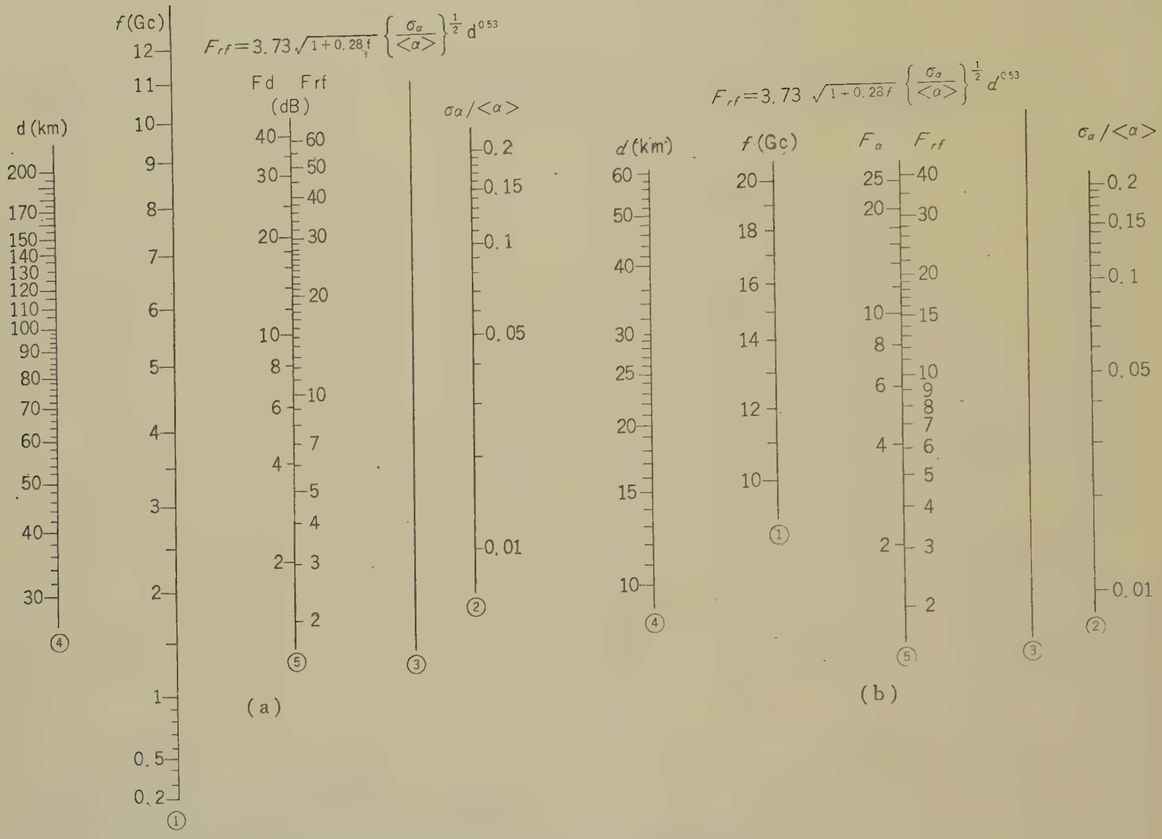

$F_{rf}=3.73\sqrt{1+0.28f}\left\{\frac{\sigma_a}{<\alpha>}\right\}^{\frac{1}{2}}d^{0.53}$
d (km)
f (Gc)
Fd Frf (dB)
σa/<α>
①
②
③
④
⑤
(a)
$F_{rf}=3.73\sqrt{1+0.28f}\left\{\frac{\sigma_a}{<\alpha>}\right\}^{\frac{1}{2}}d^{0.53}$
d (km)
f (Gc)
F_a F_{rf}
σa/<α>
(b)

Fig. 39—Nomograph $F_{rf}=3.73\sqrt{1+0.28f}\{\sigma_a/\langle\alpha\rangle\}^{1/2}d^{0.53}$.

In Japan, aerological observatories are located at intervals of about 400 km. There are not enough to meet our desire, and is not considered proper to interpolate atmospheric radio meteorological constants between observatories.

In addition, observatories are all on ground, and as a result, it is nearly impossible to find atmospheric radio meteorological constants in mountainous districts.

But it has been found that the fading range for the Mt. Futago-Tsukuba Hill path fairly distant (about 100 km) from the nearest radiosonde observatory can be explained with the atmospheric meteorological coefficients at Tateno, and that the fading range for the Toyama-Asahi path can be explained with the atmospheric coefficients at Wajima.

The Toyama-Asahi path is about 80 km away from Wajima and is a slanting path nearly parallel to the coast line. A part of it is a so-called typical propagation path along the coast-line over the sea.

Good correspondence is found between the fading range of these paths estimated from radiosonde data and the measured value. As a result, excepting mountainous districts, sonde data are considered applicable to fairly wider regional range.

Of course, because of peculiar local meteorological conditions in the lower layer, atmospheric radio meteorological constants are probably quite complicated. Therefore it is desirable to consider the meteorological conditions which cause irregular distribution of refractive index to the full.

What should be stressed here is the abnormality of fading due to abnormal meteorologi-

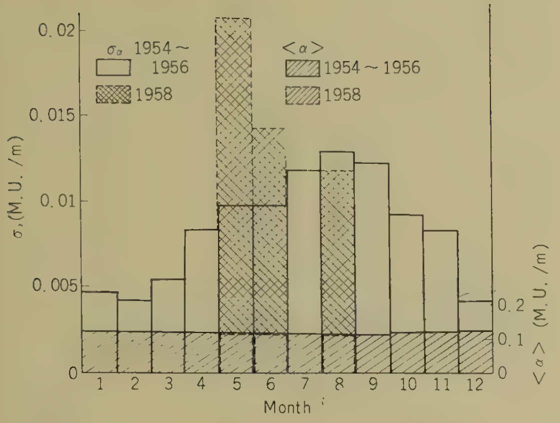


Fig. 40—Comparison between atmospheric radio meteorological constant at Wajima in 1957 and that in 3 years from 1954 to 1956.

cal condition. For instance, the fading range which occurred during the 5 months from June to October 1957 was larger than that during the summer of the year for the Fudō-tōge–Itsukaichi path; and in May and June 1958, unusually severe communication disturbances occurred often on the circuit in the northern district.

Fig. 40 shows a comparison of the atmospheric radio meteorological constants calculated from the Wajima radiosonde data in May, June, and August of 1958 with the values for the three years from 1954 to 1956. It is clearly found that the values in May and June, 1958 are unusual. (by F-test).

Of course, the appearance of the abnormal values of atmospheric radio meteorological constants is considered due to abnormal meteorological phenomena, but the reason has not yet been clarified. But the determination of the cycle with which these phenomena repeat, that is, the estimation of the return period, is a problem very important to practical radio communication.

It is considered necessary at least to calculate the atmospheric radio meteorological constants when abnormal fading appears and to compare them with those in normal years to ascertain whether there is abnormality due to propagation.

In the future, the wideband radio circuits will be expanded and the main microwave relay circuits will be loop connected; therefore will become possible to avoid disturbance due to local abnormal propagation phenomena, by selecting the proper loop-connection.

For this purpose, as in the case of shortwave propagation, also, Prediction and Alarm of troposphere propagation are believed to be necessary to make plans for avoiding the effects of abnormal meteorological conditions.

5. Conclusion

In this paper, researches on the concrete problem of propagation have been described, especially those on the estimation of receiving field variation which have been faced so far in the designing of microwave relay circuits.

This research is summed up as follows:

(1) The representation of fading range (the ratio of field intensity at cumulative probability 1% to that at cumulative probability 99%, in dB) is defined, and the correlations between various meteorological elements–e.g. vapor pressure, temperature, humidity, and modified refractive index M on the ground and mean modified refractive index gradient below 1000 m above the ground–are discussed. As a result, it is found that fading range is closely related to the variation of σ_α.

(2) There are two modes of fading, that is, long-period and short-period variation modes.

(3) In long-period variation, frequency characteristics are seldom found. In short-period variation, the frequency selective characteristics are remarkable.

(4) In the case of severe short-period variation, the fading ratio is quite dependent on frequency.

From experimental results, the causes of long- and short-period variations are thought to be different.

The former is considered due to the spatial density of received radio energy caused at the bending part of the M profile in the geometric optical approximation, while the latter is due to the mutual interference between multiple waves caused by multiple-refraction.

From the results of long-term close observation of radiosonde, the following are found.

(5) The vertical gradient α of refractive index in the layer with thickness ΔH (m) has been found to have Gaussian distribution under practical approximation. When the layer thickness is equal, the standard deviation σ_α of α has little relation to the height of the layer.

(6) Against every region and every month, the law $\Delta H \cdot \sigma_\alpha^2 =$invariant is found approximately.

(7) In this way, a method of estimating statistically the refractive index of the air has been obtained for optional regions from radiosonde data.

(8) By means of geometric optical analysis, fading range and air refractive index distribution are statistically connected quantitatively.

These results are believed to be widely applicable to not only Japan but to any region where radiosonde observation is carried out.

The author wishes to thank his collaborators in the Radio Propagation Research Section of the ECL. He is also indebted to members of the Meteorological Agencies.

Appendix

(i) Two Variables Following respectively Gamma-distribution and Beta-distribution

Suppose that x and y respectively follow Gamma-distribution function $f_0(x)$ and Beta-distribution function $f_f(y)$. Where,

$$\left.\begin{aligned} f_0(x) &= \frac{a^{\nu+1}}{\Gamma(\nu+1)} x^{\nu} e^{-ax}, \\ f_f(y) &= \frac{1}{B(n_1, n_2)} y^{n_1-1}(1-y)^{n_2-1} \end{aligned}\right\} \qquad (1)$$

If the distribution function of the independent product z is shown by $f_r(z)$, $f_r(z)$ is as follows:

$$\left.\begin{aligned} f_r(z) &= \int_0^\infty f_0(x) f_f(y) \frac{\partial(x, y)}{\partial(x, z)} dx \\ &= \frac{a^{\nu+1}}{\Gamma(\nu+1)} \cdot \frac{\Gamma(n_1+n_2)}{\Gamma(n_1)\Gamma(n_2)} \\ &\quad \times \int_z^\infty x^{\nu-1} e^{-ax} \left(\frac{z}{x}\right)^{n_1-1} \left(1-\frac{z}{x}\right)^{n_2-1} dx \\ &= \frac{a^{\nu+1}}{\Gamma(\nu+1)} \frac{\Gamma(n_1+n_2)}{\Gamma(n_1)\Gamma(n_2)} z^{n_1-1} \\ &\quad \times \int_z^\infty x^{\nu-n_1-n_2+1} e^{-ax} (x-z)^{n_2-1} dx \end{aligned}\right\} \qquad (2)$$

If $\nu+1=n_1+n_2$, Eq. (2) becomes

$$f_r(z) = \frac{a^{n_1}}{\Gamma(n_1)} z^{n_1-1} e^{-az} \qquad (3)$$

(ii) Two Variables Following Gamma-Distribution

Suppose that x and y respectively follow the Gamma-distribution functions, that is $f_0(x)$ and $F_f(y)$. If the distribution function of the independent product z is shown by $f_r(z)$, the cumulative distribution function $F_r(z)$ is obtained by the following equation:

$$\begin{aligned} F_r(z) &= \int_0^z f_r(z)\, dz \\ &= \sum_{m=1}^{n+2} C_m \int_0^\zeta \frac{\zeta^{2\nu_1+m} e^{-\zeta}}{\Gamma(2\nu_1+m+1)} d\zeta \end{aligned} \qquad (4)$$

where

$$C_m = \frac{(2\nu_1+1)\ (2\nu_1+2)(2\nu_1+3)\cdots\cdots(2\nu_1+m)}{(2\nu_1+1)\ (2\nu_1+3)(2\nu_1+5)\cdots\cdots(2\nu_1+3+2n)}$$

$$\times \frac{(2n-m+3)!\ 2^{-n+m-2}}{(m-1)!\ (n-m+2)!},$$

$$\zeta = 2\sqrt{a b z}$$

where, $$f_0(x)=\frac{a^{\nu_1+1}}{\Gamma(\nu_1+1)}x^{\nu_1}e^{-ax}$$

$$f_f(y)=\frac{b^{\nu_2+1}}{\Gamma(\nu_2+1)}y^{\nu_2}e^{-by}$$

$$\nu_2-\nu_1=n+\frac{3}{2}, \quad n: \text{integer}\geq -1$$

(iii) Two Variables respectively Following the Logarithmic Gaussian-Distribution and Gamma-Distribution

Suppose that x and y respectively follow the logarithmic Gaussian-distribution $f_0(x)$ and Gamma-distribution $f_f(y)$ as next equation:

$$\left.\begin{aligned} f_0(x) &= \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(\log_e x/x_m)^2}{2\sigma^2}}\cdot\frac{1}{x}, \\ f_f(y) &= \frac{a^{\nu+1}}{\Gamma(\nu+1)}y^{\nu}e^{-ay}. \end{aligned}\right\} \quad (5)$$

The cumulative distribution function $F_r(z)$ of the independent product $z=x\cdot y$ is as follows:

$$F_r(z)=\int_0^z dz\int_0^\infty f_0(x)f_f(y)\frac{\partial(x,y)}{\partial(x,z)}dx$$

$$=\int_0^\infty\int_0^z\frac{1}{\sigma\sqrt{2\pi}}\cdot\frac{1}{\Gamma(\nu+1)}e^{-\frac{(\log_e x/x_m)^2}{2\sigma^2}}\cdot\frac{1}{x}\left(\frac{az}{x}\right)^{\nu}e^{-\frac{az}{x}}d\left(\frac{az}{x}\right)\cdot dx$$

$$=\frac{1}{\sigma\sqrt{2\pi}}\int_0^\infty e^{-\frac{(\log_e x/x_m)^2}{2\sigma^2}}\left\{1-e^{-\frac{az}{x}}\sum_{s=0}^{\nu}\frac{\left(\frac{az}{x}\right)^s}{s!}\right\}\frac{dx}{x}.$$

where $X=\log_e\dfrac{x}{x_m}$.

$$F_r(z)=1-\frac{1}{\sigma\sqrt{2\pi}}\sum_{s=0}^{\nu}\frac{\left(\frac{az}{x_m}\right)^s}{s!}\int_{-\infty}^{\infty}\exp-\left\{\frac{X^2}{2\sigma^2}+sX+\frac{az}{x_m}e^{-X}\right\}dX.$$

where $X=\xi-(s-1)\sigma^2$

$$F_r(z)=1-\frac{1}{\sigma\sqrt{2\pi}}\sum_{s=0}^{\nu}\frac{\left(\frac{az}{x_m}\right)^s}{s!}e^{-\frac{\sigma^2}{2}+\frac{\sigma^2s^2}{2}}\int_{-\infty}^{\infty}\exp-\left\{\frac{\xi}{2\sigma^2}+\xi+\frac{az}{x_m}e^{\sigma^2(s-1)-\xi}\right\}d\xi. \quad (6)$$

The integral term of Eq. (6) is K. W. Wagner's after-effect function.

That is $$F(p,q)=\int_{-\infty}^{\infty}e^{-q^2z-z-pe^{-z}}dz.$$

Accordingly, by Wagner's expression, Eq. (6) is shown as

$$F_r(z)=1-\frac{e^{-\frac{\sigma^2}{2}}}{\sigma\sqrt{2\pi}}\sum_{s=0}^{\nu}\frac{\left(\frac{az}{x_m}\right)^s}{s!}e^{\frac{\sigma^2s^2}{2}}f\left(\frac{az}{x_m}e^{\sigma^2(s-1)},\frac{1}{\sigma\sqrt{2}}\right).$$

References

(1) For example, C. R. Barrows and S. S. Attwood: "Radio Wave Propagation", Consolidated Summary Technical Report of the Committee on Propagation of the National Defence Research Committee, New York, 1949.

(2) D. E. Kerr: Propagation of Short Radio Waves." *Radiation Laboratory Series*, Vol. 13,

Mc Graw Hill, 1951.

(3) Saburo MATSUO, Shigetaka UGAI, Kiyoshi KAKITA, Fumio IKEGAMI and Yoshinori KONO: "Microwave Fading", *Report of E. C.L.*, Vol. 1, No. 3, pp 38~47, Nov., 1953.

(4) For example, Radio Wave Propagation Consolidated Summary Technical Reports, pp. 168–172, pp. 183–190, 1958.

(5) Ken-ichi MAEDA and Mitsuo GOTO: Radio-Wave Propagation, Iwanami Book Series, No. 185, 1958. (in Japanese).

(6) Shigetaka UGAI: Statistical Consideration of the Structure of Atmospheric Refractive Index, *Report of E.C.L.*, Vol. 7, No. 8, pp. 253–290, Aug., 1959.

(7) Jahnke and Emde: Tables of Functions, p. 112, 1933.

* * * *

REVIEW OF THE ELECTRICAL COMMUNICATION LABORATORY
VOLUME 9, NUMBERS 5-6, MAY-JUNE, 1961

U.D.C. 621.391.827.42:621.372:621.372.823

Interference Waves in Waveguide Transmission and Intermodulation Noise in FM System for Waveguide Communication*

Kazuhiro MIYAUCHI†

In the circular waveguide transmission line for long distance communication, echo distortion is caused by interference waves which are produced by mode conversion-reconversion and multiple reflection. This paper deals theoretically with the characteristics of the interference waves and the intermodulation noise caused by these waves in a frequency-modulation system for the transmission of a frequency-division multiplex telephone signal. The effects of system design parameters and waveguide parameters on the intermodulation noise level are also examined.

1. Introduction

In this paper some characteristics of frequency modulation communication through waveguides are described.

It is considered that there are two possible modulation methods—pulse code and frequency modulation—for waveguide communication. It has not yet been concluded which of these two modulation methods is better. Pulse code modulation is undoubtedly superior for very long distance communication, because it keeps out noise accumulation in the long chains of repeaters and therefore does not require highly precise waveguides. However, there is the possibility of applying frequency modulation to medium-long-distance communication because of the simplicity of the terminal and repeating devices.

The ability to manufacture suitably precise waveguides at reasonable cost is required for the construction of frequency-modulation systems, but at present it has not been quantitatively determined how precise the waveguides must be. One purpose of this paper is to report an estimation of the precision necessary for waveguides which are used for frequency modulation transmission on a frequency-division multiplex telephone system.

Larsen(1) discussed this problem for some simple cases. But the results of his treatment cannot be applied to the case of large modulation index which is considered to be practical, and the presence of spurious modes except TE_{02} is not considered correctly, moreover the insertion of mode filters is not discussed.

In this paper, general approximation formulas for echo distortion, which can be applied to the case in which any spurious mode is present and any modulation index is used, are deviced. Approximation formulas for waveguide lines with mode filters are also derived.

In the course of deriving the distortion formulas, it is necessary to examine the characteristics of interference waves produced by mode conversion and reconversion phe-

---

* MS received by the Electrical Communication Laboratory, on October 3, 1960. Originally published in the *Kenkyû Zituyôka Hôkoku* (*Electrical Communication Laboratory Technical Journal*), Vol. 10, No. 4, pp. 587-627, April, 1961.

† Hyper-frequency Research Section.

nomena in multimode waveguide lines. Some properties of the interference waves are discussed in Section 2. In the case of the transmission lines with two propagating modes it is possible to treat the interference waves fairly correctly. For that case, the power sum law relationship between the interference waves is examined, and an approximate error is discussed for the case when only the interference waves converted twice are considered, and then the mean delayed power spectrum of the interference waves is derived.

The above-mentioned results have been derived for a two-mode transmission line with discrete mode coupling, but they are approximately generalized to the case of many modes with continuous mode coupling. The condition for the generalization is considered approximately.

After these treatments, the total powers of the interference waves are calculated, and echo power due to signal mode reflection is compared with the power of interference waves due to mode conversion and reconversion. The mean delayed power spectrum of the interference waves in waveguide lines with mode filters is calculated.

In Section 3, the results obtained in Section 2 are applied to calculate the intermodulation noise appearing in the multiplex telephone signal due to the interference waves. In this calculation the echo distortion formulas given by Albersheim,⁽²⁾ Scafer, Medhurst,⁽³⁾ and Small are multiplied by the mean delayed power spectrum of the interference waves, and then integrated with respect to the delay time.

Using these intermodulation noise (or echo distortion) formulas, the effects of spurious mode attenuation and frequency deviation on the intermodulation noise level are examined. To calculate the waveguide tolerances allowable for frequency-modulation communication through multimode waveguides, the relation between the waveguide deformation and mode conversion must be known, and allowable maximum intermodulation noise level for practical communication systems must be assumed.

The values of the mode conversion coefficients at the junctions of circular waveguides have been calculated in many papers,⁽¹⁾,⁽⁴⁾,* but little is known about continuous conversion. Therefore in this paper the calculations are restricted to discrete mode coupling at the waveguide junctions.

As to the allowable intermodulation noise level, practical values of system design parameters are assumed, and the noise level which satisfies the recommended standard of the C.C.I.R. is calculated.

After these preparatory calculations, allowable tolerances for copper-pipe waveguide lines and helix waveguide lines are calculated.

Calculations for waveguide lines with mode filters are more complicated, but the methods used for uniform waveguide lines are applicable. Intermodulation noise formulas are calculated and allowable tolerances for helix waveguide lines with mode filters are also calculated.

2. Characteristics of Interference Waves

2.1. Interference Waves in a Transmission Line Having Two Propagating Modes

Here we will consider a straight transmission line made of many waveguides having random mode conversion. It is assumed that there are only two propagating modes.

N waveguides of nearly equal length are connected in cascade, and mode conversion is assumed to be present only at the $N+1$ junctions including the junctions to the input and output mode excitors.

The length of each waveguide is $l_1, l_2, \cdots, l_N$ from input to output. We assign the numbers $0, 1, 2, \cdots, N$ to each junction point in the same order. The propagation constant of the signal mode TX_1 is γ_1, and the propagation constant of the spurious mode TX_2 is γ_2. We let S_i and k_i be mode coupling coefficients between the same modes and between the different modes respectively at the i-th junction.

If the influence of the reflection at the junction is neglected, signal amplitude of the TX_1 mode at the output is represented by the summation of $V(p_0, p_1, \cdots\cdots, p_N : q_1, q_2, \cdots\cdots, q_N)$ of equation (1)

---

* Noda: unpublished paper.

$$\begin{aligned} &V(p_0, p_1, \cdots\cdots, p_N ; q_1, q_2, \cdots\cdots, q_N) \\ &= S_0^{(1-p_0)} \cdot S_1^{(1-p_1)} \cdot \cdots\cdots \cdot S_N^{(1-p_N)} \cdot k_0^{p_0} \cdot k_1^{p_1} \\ &\cdot \cdots\cdots \cdot k_N^{p_N} \cdot e^{-\gamma_1 l_1 (1-q_1)} \cdot e^{-\gamma_1 l_2 (1-q_2)} \cdot \cdots\cdots \\ &\cdot e^{-\gamma_1 l_N (1-q_N)} \cdot e^{-\gamma_2 l_1 q_1} \cdot e^{-\gamma_2 l_2 q_2} \cdot \cdots\cdots \cdot e^{-\gamma_2 l_N q_N}. \end{aligned} \tag{1}$$

Where p_i and q_i are parameters which take the value of 0 or 1. Parameter p_i is set to 0 if V of equation (1) corresponds to the received signal which is converted from one mode to another at the i-th junction, and set to 1 if V corresponds to the received signal transmitted with the mode unaltered through the i-th junction. Parameter q_i is set to 0 if V corresponds to the received signal which is transmitted with the TX_1 mode through the i-th waveguide, and set to 1 if V corresponds to the received signal which is transmitted with the TX_2 mode through the i-th waveguide.

In order that equation (1) represents the amplitude of the TX_1 signal mode at the output of the waveguide line, two condition must be satisfied:

$$p_0 + p_1 + \cdots\cdots + p_N = \text{an even integer}, \tag{2}$$

and

$$\begin{aligned} p_i &= q_i(1-q_{i+1}) + (1-q_i)b_{i+1} \\ &= (q_i + q_{i+1}) - 2q_i q_{i+1} \\ &\qquad i = 0, 1, 2, \cdots\cdots, N \end{aligned} \tag{3}$$

where q_0 and q_{N+1} are defined as zero.

In this paper the expression "main wave" means V, in equation (1), which is transmitted through all waveguides with the TX_1 mode; and "interference wave" means V, in equation (1), which is transmitted through more than one waveguide with the TX_2 mode.

We define two variables n and m for the interference wave: n is the number of waveguides in which the interference wave is transmitted with the TX_2 mode, and m represents the number of the mode conversion.

Representation corresponding to the main wave is

$$\begin{aligned} &V(0, 0, \cdots\cdots, 0 ; 0, 0, \cdots\cdots, 0) \\ &= S_0 \cdot S_1 \cdot \cdots\cdots \cdot S_N \cdot e^{-\gamma_1 l_1} \cdot e^{-\gamma_1 l_2} \cdot \cdots\cdots \cdot e^{-\gamma_1 l_N} \\ &= S \cdot e^{-\alpha_1 L} \cdot e^{-j\beta_1 L} \end{aligned} \tag{4}$$

where

$$\gamma_1 = \alpha_1 + j\beta_1 \tag{5}$$

$$S = S_0 \cdot S_1 \cdot \cdots\cdots \cdot S_N \tag{6}$$

$$L = l_1 + l_2 + \cdots\cdots + l_N. \tag{7}$$

α_1 and β_1 are the attenuation and phase constant of TX_1 mode. S is the main wave attenuation caused by the excitation of the unwanted TX_2 mode. L is the total length of the line composed of N unit waveguides (l_i in length).

We let $W(p_0, p_1, \cdots\cdots, p_N ; q_1, q_2, \cdots\cdots, q_N)$ be the amplitude ratio of the interference wave to the main wave.

$$\begin{aligned} &W(p_0, p_1, \cdots\cdots, p_N ; q_1, q_2, \cdots\cdots q_N) \\ &= \frac{V(p_0, p_1, \cdots\cdots, p_N ; q_1, q_2, \cdots\cdots, q_N)}{V(0, 0, \cdots\cdots, 0 ; 0, 0, \cdots\cdots, 0)} \\ &= \left(\frac{k_0}{S_0}\right)^{p_0} \cdot \left(\frac{k_1}{S_1}\right)^{p_1} \cdot \cdots\cdots \cdot \left(\frac{k_N}{S_N}\right)^{p_N} \\ &\cdot e^{-\Delta\gamma l_1 q_1} e^{-\Delta\gamma l_2 q_2} \cdot \cdots\cdots \cdot e^{-\Delta\gamma l_N q_N} \end{aligned} \tag{8}$$

where

$$\Delta\gamma = \gamma_2 - \gamma_1.$$

In the waveguides of good mechanical accuracy S_i is considered to be almost equal to 1. Equation (8) is rewritten as:

$$\begin{aligned} &W(p_0, p_1, \cdots\cdots, p_N ; q_1, q_2, \cdots\cdots, q_N) \\ &= k_0^{p_0} \cdot k_1^{p_1} \cdot \cdots\cdots \cdot k_N^{p_N} \cdot e^{-\Delta\gamma(l_1 q_1 + l_2 q_2 + \cdots\cdots + l_N q_N)} \end{aligned} \tag{9}$$

Sometimes W is called the "interference wave" in place of V in the present paper.

Using $W(p_0, p_1, \cdots\cdots, p_N ; q_1, q_2, \cdots\cdots, q_N)$, the amplitude of the TX_1 mode received at the output of the waveguide line is represented as:

$$T=Se^{-\gamma_1 L}[1+\sum W(p_0, p_1, \cdots\cdots, p_N ; q_1, q_2, \cdots\cdots, q_N)] \tag{10}$$

where the summation extends over the combination of $p_0, p_1, \cdots\cdots, p_N ; q_1, q_2, \cdots\cdots, q_N$, so long as they satisfy both equations (2) and (3).

We define some partial sums of equation (10) as follows:

$$\hat{U}(N, n, m)=\sum_{\substack{p_0+p_1+\cdots\cdots+p_N=\text{even number } m \\ q_1+q_2+\cdots\cdots+q_N= \quad n}} W(p_0, p_1, \cdots\cdots, p_N ; q_1, q_2, \cdots\cdots q_N) \tag{11}$$

$$\hat{U}(N, n)=\sum_{m=2,4,6,\cdots\cdots,2\left[\frac{N+1}{2}\right]} U(N, n, m) \tag{12}$$

where $[x]$ means the largest integer which does not exceed x.

We let the number of terms in the right-hand side of equation (11) be $C(N, n, m)$, which is shown in Appendix 2 to be as follows:

$$C(N, n, m)={}_{n-1}C_{(m/2)-1}\cdot{}_{N-n+1}C_{m/2}. \tag{13}$$

Where we assume that ${}_0C_0$ is equal to 1, and the other ${}_pC_q$'s having no meaning are equal to zero.

We will consider l_i is slightly different from the standard length l.

$$\Delta l_i=l_i-l. \tag{14}$$

Equation (15) is rewritten as:

$$\begin{aligned} &W(p_0, p_1, \cdots\cdots, p_N ; q_1, q_2, \cdots\cdots, q_N) \\ &=k_0{}^{p_0}\cdot k_1{}^{p_1}\cdot\cdots\cdots\cdot k_N{}^{p_N}\cdot e^{-\Delta\gamma nl} \\ &\cdot e^{-\Delta\gamma(\Delta l_1 q_1+\Delta l_2 q_2+\cdots\cdots+\Delta l_N q_N)}. \end{aligned} \tag{15}$$

$\Delta\gamma$ is the difference of propagation constants of the two modes TX_2 and TX_1.

$$\Delta\gamma=\Delta\alpha+j\Delta\beta \tag{16}$$

$$\Delta\alpha=\alpha_2-\alpha_1 \tag{17}$$

$$\Delta\beta=\beta_2-\beta_1. \tag{18}$$

Usually Δl_i is very small and $\Delta\alpha$ is much smaller than $\Delta\beta$, so the exponent of the last factor in equation (15) is rewritten as:

$$\begin{aligned} &-\Delta\gamma(\Delta l_1 q_1+\Delta l_2 q_2+\cdots\cdots+\Delta l_N q_N) \\ &=-j\Delta\beta(\Delta l_1 q_1+\Delta l_2 q_2+\cdots\cdots+\Delta l_N p_N). \end{aligned} \tag{19}$$

Accordingly, we get:

$$\begin{aligned} &W(p_0, p_1, \cdots\cdots p_N ; q_1, q_2, \cdots\cdots, q_N) \\ &=k_0{}^{p_0}\cdot k_1{}^{p_1}\cdot\cdots\cdots\cdot k_N{}^{p_N}\cdot e^{-\gamma nl} \\ &\cdot e^{-j\Delta\beta(\Delta l_1 q_1+\Delta l_2 q_2+\cdots\cdots+\Delta l_N q_N)}. \end{aligned} \tag{20}$$

We will separate the factor $e^{-\Delta\gamma nl}$ from $W(p_0, p_1, \cdots\cdots, p_N ; q_1, q_2, \cdots\cdots, q_N)$, $U(N, n, m)$, and $\hat{U}(N, n)$.

$$\begin{aligned} &W(p_0, p_1, p_N ; q_1, q_2, \cdots\cdots, q_N) \\ &=e^{-\Delta\gamma nl}w(p_0, p_1, \cdots\cdots, p_N ; q_1, q_2, \cdots\cdots, q_N) \end{aligned} \tag{21}$$

$$U(N, n, m)=e^{-\Delta\gamma nl}u(N, n, m) \tag{22}$$

$$\hat{U}(N, n)=e^{-\Delta\gamma nl}\hat{u}(N, n). \tag{23}$$

That is

$$\begin{aligned} &w(p_0, p_1, \cdots\cdots, p_N ; q_1, q_2, \cdots\cdots, q_N) \\ &=k_0{}^{p_0}\cdot k_1{}^{p_1}\cdot\cdots\cdots\cdot k_N{}^{p_N} \\ &\cdot e^{-j\Delta\beta(\Delta l_1 q_1+\Delta l_2 q_2+\cdots\cdots+\Delta l_N q_N)} \end{aligned} \tag{24}$$

$$\hat{u}(N, n, m) = \sum_{\substack{p_0+p_1+\cdots\cdots+p_N = \text{even number } m \\ q_1+q_2+\cdots\cdots+q_N = n}} w(p_0, p_1, \cdots\cdots, p_N\,; q_1, q_2, \cdots\cdots, q_N) \tag{25}$$

$$\hat{u}(N, n) = \sum_{m=2,4,6,\cdots\cdots,2\left[\frac{N+1}{2}\right]} u(N, n, m). \tag{26}$$

It may be considered that $w(p_0, p_1, \cdots\cdots, p_N; q_1, q_2, \cdots\cdots q_N)$, $u(N, n, m)$, and $\hat{u}(N, n)$ vary only slightly in the narrow frequency band which is used for the transmission of a communication signal. Therefore, if we rewrite equation (10) as equation (27), the frequency characteristics of each term in the later are represented only by $e^{-\Delta\gamma nl}$.

$$T = Se^{-\gamma_1 L}\left[1 + \sum_{n=1}^{N} e^{-\Delta\gamma nl}\hat{u}(N, n)\right]$$

$$= Se^{-\gamma_1 L}\left[1 + \sum_{n=1}^{N} e^{-j\Delta\beta nl}e^{-\Delta\alpha nl}\hat{u}(N, n)\right]. \tag{27}$$

Each term of the summation in equation (27) represents the amplitude ratio and phase difference of the interference wave to the main wave.

Now we will discuss the frequency characteristics of the interference wave in equation (27). Expanding $\Delta\beta nl$ in a Taylor series around the center angular frequency ω_0 of the signal band, we obtain

$$\Delta\beta nl = [\Delta\beta]_{\omega_0} nl + \left[\frac{\partial(\Delta\beta)}{\partial\omega}\right]_{\omega_0} nl(\omega-\omega_0)$$

$$+ \frac{1}{2}\left[\frac{\partial^2(\Delta\beta)}{\partial\omega^2}\right]_{\omega_0} nl(\omega-\omega_0)^2 + \cdots\cdots. \tag{28}$$

Equation (28) represents the change of signal form which has been carried by the interference wave in the transmission through the multimode waveguides; the first term corresponds to the phase shift of the carrier, the second term corresponds to the group delay, and the others are factors contributing to wave form distortion.

There are many studies concerning the effects of the phase shift and the group delay of the interference wave on the signal distortion when the interference wave is received with the main wave. On the other hand, the effect of the waveform distortion of the interference wave on the signal distortion has not been treated directly. In practical treatments the influence of the waveform distortion of the interference wave is usually neglected. Therefore in this paper we will represent the characteristics of the interference wave only with the phase shift and group delay.

The phase shift of the carrier consists of

$$\Delta\beta(\Delta l_1 q_1 + \Delta l_2 q_2 + \cdots\cdots + l_N q_N)$$

in equation (20) or (24), and of the first term in the right-hand side of equation (28).

Group delay time t is given as:

$$t = \frac{\partial(\Delta\beta)}{\partial\omega} nl = nl\left(\frac{\partial\beta_2}{\partial\omega} - \frac{\partial\beta_1}{\partial\omega}\right)$$

$$= nl\left(\frac{1}{v_{g_2}} - \frac{1}{v_{g_1}}\right) = nl(\tau_2 - \tau_\lambda) \tag{29}$$

where

v_{g_1}, v_{g_2} = group velocity of the TX_1 and TX_2 modes

τ_1, τ_2 = group delay time of the TX_1 and TX_2 modes per unit length respectively.

Now we consider the average characteristics of the interference waves for an ensemble of waveguide lines having statistically identical constructions. We assume that k_i and Δl_i in equation (24) are all random variables, and their statistical characters are defined as:

1. k_i and Δl_i are all independent random variables.
2. All k_i's have the same probability distribution; mean value is zero, and standard deviation is σ_K.
3. All Δl_i's have the same probability distribution, and mean value is zero.

We will define the covariance between two different interference waves as follows;

$$\hat{C}[W(p_0, p_1, \cdots\cdots, p_N; q_1, q_2, \cdots\cdots, q_N)\cdot W(p_0', p_1', \cdots\cdots, p_N'; q_1', q_2', \cdots\cdots, q_N')]$$
$$=\overline{W(p_0, p_1, \cdots\cdots, p_N; q_1, q_2, \cdots\cdots, q_N)\cdot W^*(p_0', p_1', \cdots\cdots, p_N'; q_1', q_2', \cdots\cdots, q_N')}. \qquad (30)$$

Where the asterisk and bar indicate complex conjugate and ensemble average respectively.

The average power of the interference wave is identical with its amplitude variance, which is defined as follows:

$$P[W(p_0, p_1, \cdots\cdots, p_N; q_1, q_2, \cdots\cdots, q_N)]$$
$$=\overline{W(p_0, p_1, \cdots\cdots, p_N; q_1, q_2, \cdots\cdots, q_N)\cdot W^*(p_0, p_1. \cdots\cdots, p_N; q_1, q_2, \cdots\cdots, q_N)}. \qquad (31)$$

It is easily seen that the covariance between two different interference waves reduces to zero, if we take account of assumptions (1), (2), (3), and equation (24).

Using these results we can derived the power sum law for the interference waves:

$$\begin{aligned}&P[W(p_0, p_1, \cdots\cdots, p_N; q_1, q_2, \cdots\cdots, q_N)\\&\quad+W(p_0', p_1', \cdots\cdots, p_N'; q_1', q_2', \cdots\cdots, q_N')\\&\quad+W(p_0'', p_1'', \cdots\cdots, p_N''; q_1'', q_2'', \cdots\cdots, q_N'')\\&\quad+\cdots\cdots]\\&=P[W(p_0, p_1, \cdots\cdots, p_N; q_1, q_2, \cdots\cdots, q_N)]\\&\quad+P[W(p_0', p_1', \cdots\cdots, p_N'; q_1', q_2, \cdots\cdots, q_N')]\\&\quad+P[W(p_0'', p_1'', \cdots\cdots, p_N''; q_1'', q_2'', \cdots\cdots, q_N'')\\&\quad+\cdots\cdots].\end{aligned} \qquad (32)$$

The average power of w, u, and $\hat{u}$ indicated in equations (24), (25), and (26) are represented as:

$$P[w(p_0, p_1, \cdots\cdots, p_N; q_1, q_2, \cdots\cdots, q_N)]=\sigma_K{}^{2m} \qquad (33)$$

$$P[u(N, n, m)]=\sigma_K{}^{2m}C(N, n, m) \qquad (34)$$

$$P[\hat{u}(N, n)]=\sum_{m=2,4,6,\cdots\cdots,2[\frac{N+1}{2}]} P[u(N, n, m)]. \qquad (35)$$

From equation (23) and (29) the total average power of the interference waves having the group delay time of $nl(\tau_2-\tau_1)$ becomes $P[\hat{u}(N, n)]\cdot e^{-2\Delta\alpha nl}$.

Each term in the summation of equation (27) can be considered to represent the relation between the average power and the group delay time of the interference wave. This relation is sketched in Fig. 1.

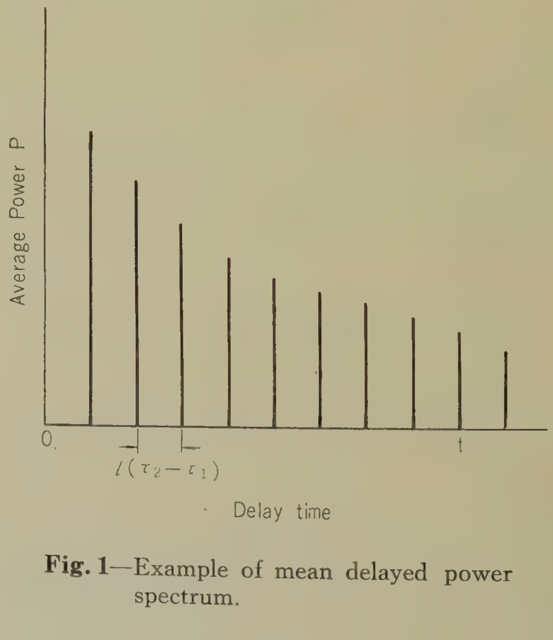


Fig. 1—Example of mean delayed power spectrum.

In this paper we call this relation the "mean delayed power spectrum" of the interference waves.

In the practical treatment it is convenient

---

* Standard deviation σ_K is defined as:

$$\sigma_K{}^2=\overline{k_i\cdot k_i{}^*}.$$

Where the asterisk means complex conjugate, and the bar means ensemble average.

to define the mean power density of the interference waves per unit delay time in order to compare the interference waves caused by different spurious modes.

$$p=\frac{P[\hat{u}(N,n)]}{l\ \tau_2-\tau_1}e^{-2\Delta\alpha n l}. \qquad (36)$$

Though $P[\hat{u}(N,n)]$ consists of the number of interference wave components, it is possible to derive its approximation formula, if a condition is satisfied between σ_K and N. We will derive this approximating condition here.

Comparing the successive two terms in equation (26), we obtain.

$$\frac{P[u(N,n,2r)]}{P[u(N,n,2r-2)]}$$

$$=\frac{(n-r+1)(N-n-r+2)}{r(r-1)}\sigma_K{}^4 \qquad (37)$$

where we have $m=2\gamma$, because m is an even integer.

It is assumed that $N\gg1$; therefore the maximum value of equation (37) with respect to n and γ is approximately $1/8\ N^2\sigma_K{}^4$. Therefore, if this value is sufficiently small it is possible that $P[\hat{u}(N,u)]$ is represented only by the first term. That is,

$$P[\hat{u}(N,n)]=\sigma_K{}^4(N-n+1), \quad \sigma_K{}^4\ll\frac{8}{N^2}. \qquad (38)$$

Equation (38) means that if the mode conversion coefficient at the junction is sufficiently small we can neglect the interference waves which are converted more than twice.

Using equation (38) we rewrite equation (36) to the representation for the average power density having the delay time from t to $t+dt$. The result is shown as:

$$p(t)=\frac{\sigma_K{}^4|\Gamma-t|}{l^2(\tau_2-\tau_1)^2}e^{-2\Delta\alpha\frac{t}{\tau_2-\tau_1}}, \quad 0\leq\frac{t}{\tau_2-\tau_1}\leq L \qquad (39)$$

where

$$\Gamma=(N+1)l(\tau_2-\tau_1). \qquad (40)$$

Γ equals approximately the delay time difference between the TX_2 and TX_1 modes for the total length of the waveguide line.

The situation discussed here has been for the case of a two-mode transmission line in which the spurious mode is single and not degenerated. Practically this situation corresponds to helix waveguide transmission lines in which the interference wave due to the TE_{02} spurious mode is dominant. On the other hand, in copper-pipe waveguide lines, the interference wave due to the TE_{12} mode is dominant, and the TE_{1n} mode has even-odd degeneration in the circular waveguide; therefore, this must be treated as another case.

Treating this situation, we can assume that there is no mode conversion between the spurious modes. If the independence between the coupling coefficients from the TE_{01} to the even and odd spurious modes are assumed together with assumptions (1), (2), and (3), we can derive the same formulas as in the the non-degenerating case. $C(N,n,m)$ is modified as:

$$C(N,n,m)=2^{m/2}\cdot{}_{n-1}C_{(m/2)-1}\cdot{}_{N-n+1}C_{m/2}. \qquad (41)$$

Approximate variance of $\hat{U}(N,n)$ becomes

$$P[\hat{u}(N,n)]=2\sigma_K{}^4(N-n+1), \quad \sigma_K{}^4\ll\frac{4}{N^2}. \qquad (42)$$

Mean delay power spectrum becomes

$$p(t)=\frac{2\sigma_K{}^4|\Gamma-t|}{l^2(\tau_2-\tau_1)^2}e^{-2\Delta\alpha\frac{t}{\tau_2-\tau_1}}, \quad 0\leq\frac{t}{\tau_2-\tau_1}\leq L. \qquad (43)$$

Now σ_K used here is the standard deviation of the coupling coefficient to the even or odd spurious mode (i.e., the standard deviation of the rectangular component of coupling coefficient vector), but as it is usual that the standard deviation of the amplitude of the coupling coefficient vector is known, we will derive the formulas using it instead of σ_K.

We let the amplitude of the coupling coefficient vector be K, and its polarization angle be θ. We will calculate the standard deviation of the rectangular component of the coupling coefficient vector assuming that K and θ are independent random variables and that the distribution of θ is uniform.

$$P(K_x) = \overline{(K\cos\theta)^2} = \frac{1}{2}P(K) \qquad (44)$$

$$P(K_y) = \overline{(K\sin\theta)^2} = \frac{1}{2}P(K) \qquad (45)$$

where $P[x]$ indicates the variance of x, and the bar indicates the ensemble average.

$P[K_x]$ and $P[K_y]$ are the same and equal to $\sigma_K{}^2$; $P[K]$ is the amplitude variance of the coupling coefficient vector. Setting $P(K]=\sigma_{K,d}{}^2$, we obtain:

$$2\sigma_K{}^2 = \sigma_{K,d}{}^2. \qquad (46)$$

Using equation (46), we can rewrite equations (42) and (43) as:

$$P[\hat{u}(n, m)] = \frac{1}{2}\sigma_{K,d}{}^4(N-n+1) \qquad (42)'$$

$$p(t) = \frac{1}{2}\ \frac{\sigma_{K,d}{}^4|\Gamma-t|}{l^2(\tau_2-\tau_1)^2}e^{-2\Delta\alpha\frac{t}{\tau_2-\tau_1}}, \qquad (43)'$$

$$0 \leq \frac{t}{\tau_2-\tau_1} \leq L.$$

2.2. Generalization of the Results of Section 2.1.

In Section 2.1. we treated waveguide transmission lines with two propagating modes and derived the power sum law between the interference waves, and discussed the approximation error which must be taken into account when we neglect the interference waves converted more than twice.

In practical circular waveguides there are many spurious modes, and their treatment is not easy. Therefore, we will assume the power sum law between the interference waves and will discuss a simple model of multimode waveguides to get a rough estimation of the approximation error which must be taken into account when we neglect the interference waves converted more than twice.

There are x propagating modes including the signal mode. Mode conversion is present at $N+1$ junctions between these x propagating modes. We assume that the conversion coefficients are all independent random variables having zero mean value and the same standard deviation σ. The mode attenuation constant is neglected in order to simplify the treatment.

We let the variance of the interference wave converted m times be P_m. Then we obtain:

$$P_m = {}_{N+1}C_m \cdot (x-1)^{m-1} \cdot \sigma^{2m}, \quad m>2. \qquad (47)$$

Comparing P_m and P_{m-1} for $m>2$, we obtain:

$$\frac{P_m}{P_{m-1}} = \frac{(N+2-m)}{m}(x-1)\sigma^2 \leq \frac{N(x-1)}{3}\sigma^2. \qquad (48)$$

If we let the average total power converted from one mode to all the other modes at a junction be ΔP; then ΔP equals $(x-1)\sigma^2$. Therefore, the right-hand side of equation (48) is rewritten as:

$$\frac{N(x-1)\sigma^2}{3} = \frac{1}{3}N\Delta P. \qquad (49)$$

Though the value shown in equation (49) is derived for a simple model of a multimode waveguide line, it may be used to furnish information about the approximation error when we take only the twice converted interference waves. If we calculate equation (49) for the practical case, its value is observed to be so small that we may conclude that the approximation with only the twice converted interference waves is also applicable to the case of many spurious modes.

Thus far, we have treated multimode transmission lines with discrete mode conversion. The following discussion is concerned with the case of continuous mode conversion. As the treatment of multimode transmission lines with continuous mode conversion is difficult if we consider the general case, we will make the problem easier by means of the following assumptions, and will derive the same formulas as in the case with discrete conversion.

We will consider a waveguide attached to its right-hand or left-hand junction as a unit element of the transmission line. We will assume that the mode conversion coefficients from the TE_{01} mode to the spurious modes in these unit elements are independent random variables having zero as the mean value. And for straight waveguide lines with discrete and/or continuous mode conversion we make the following assumptions:

(1) Covariance between two different interference wave is zero.

(2) Total power of the interference waves converted more than twice is negligibly small.

For these assumptions the characteristics of the interference waves are treated with the same method as used in the case with only discrete mode conversion. The mean delayed power spectrum becomes continuous in transmission lines with only continuous mode conversion, but becomes a mixture of discrete and continuous spectrums for the case of not only discrete but also continuous conversion. Representation of the mean delayed power spectrum as indicated in equation (43) is available for all cases with the assumption that σ_K must be considered as the standard deviation of the mode conversion coefficient per unit element of the transmission line.

2.3. Total Power of the Interference Waves

As a twice converted interference wave is concerned with only one spurious mode and the power sum law between the interference waves is assumed, the total power of all interference waves received at the end of the transmission line is represented as the sum of the power of every set of interference waves which belong to every spurious mode.

The power of a set of interference waves which belong to a spurious mode having no degeneration is represented as:

$$P = \sum_{n=1}^{N} \sigma_K{}^4 (N-n+1) e^{-2\Delta\alpha n l} \qquad (50)$$

from equation (42).

We will assume:

$$N \gg 1 \qquad (51)$$

$$2\Delta\alpha l \ll 1, \quad 2\Delta\alpha(N+1)l \gg 1. \qquad (52)$$

Then equation (51) is rewriten as:

$$P = \frac{\sigma_K{}^4 N^2}{2A} \qquad (53)$$

where

$$A = \Delta\alpha L \quad \text{(Naper)}. \qquad (54)$$

Calculating for the spurious mode having even-odd degeneration we obtain

$$P = \frac{\sigma_{K,d}{}^4 \cdot N^2}{4A}. \qquad (55)$$

It must be noticed that P represents the average total power of the interference waves with respect to unit power of the main wave.

2.4. The Influence of Echo due to Signal Mode Reflections

Up to this point the influence of mode reflection at the junctions of waveguides has been neglected. The reason is as follows: The spurious modes which produce interference waves of considerable power level are of comparatively low order. The ratio of transmitting coupling coefficient to reflecting coupling coefficient from the TE_{01} mode to these low order spurious modes is greater than about twenty in practical waveguides.* Therefore, if we consider the interference wave due to transmitting coupling coefficients is a small quantity of the first order, the interference wave due to reflecting coupling coefficients (i.e. echo wave) can be considered to be a small quantity of the second order.

But here is an exception. The level of the interference waves produced by multiple reflections of the signal mode (TE_{01} mode) is considered to be a small quantity of the first order. Therefore, we must examine their power level to determine whether they can be neglected or not.*

---

* For example, waveguide diameter 51 mm, operating frequency 50 Gc.

We will treat only those interference waves (i.e. echo waves) reflected twice.

In place of equation (27), we obtain

$$T=Se^{-\gamma L}[1+\sum e^{-2\gamma nl}\hat{u}(N, n)] \qquad (56)$$

where the variance of $\hat{U}(N, n)$ is $(N-n+1)\sigma_R{}^4$, the covariance between two different $\hat{U}(N, n)$ is zero, and σ_R is the standard deviation of the reflection coefficient of the signal mode at the junction of the waveguides.

Delay time of the interference wave is represented as:

$$t=2\frac{\partial\beta}{\partial\omega}nl=2\frac{nl}{v_g}=2nl\tau \qquad (57)$$

where

β=phase constant of the signal mode
l=length of unit waveguide
v_g=group velocity of the signal mode
τ=group delay time of the signal mode per unit length of waveguide.

Mean delayed power spectrum is

$$p(t)=\frac{\sigma_R{}^4|\Gamma-t|}{4l^2\tau^2}e^{-2\alpha\frac{t}{\tau}} \qquad (58)$$

and total power of the interference waves is

$$P=\frac{\sigma_R{}^4\cdot N^2}{2A} \qquad (59)$$

where

α=attenuation constant of the signal mode
$\Gamma=2L\tau$
$A=2L\alpha$
L=total length of the waveguide line.

We will calculate the value of equation (59) for a practical case and will compare it with equation (53) in the case of interference waves caused by TE_{01}-TE_{02} conversion-reconversion.

$L=20$ km ····total length of waveguide line
$N=4000$ ······the number of unit waveguides
$f_0=50$ Gc ····carrier frequency
$2a=51$ mm ····diameter of waveguide
$\sigma_\delta=0.1$ mm····standard deviation of the radius difference at the junction of waveguides.

Referring to Appendix 6.1., we obtain

P (reflection) ····-114 dB from equation (59)
P (conversion)····$-$ 27 dB from equation (53).

Where P (reflection) indicates the total power of the interference waves caused by TE_{01} mode reflections, and P (conversion) indicates the total power of the interference waves caused by TE_{01}-TE_{02} conversion-reconversion.

As P (reflection) is very much smaller than P (conversion), we can neglect the interference waves due to signal mode reflection in the practical treatment.

2.5. Effects of Mode Filters on the Mean Delayed Power Spectrum

We will consider the characteristics of interference waves in a waveguide line with mode filters. We let A_M and l_M be the attenuation and spacing respectively of the mode filters. Assuming the power sum law between the interference waves, we obtain the mean delayed power spectrum:

$$p(t)=\frac{\sigma_K{}^4|\Gamma-t|}{l^2\Delta\tau^2}\left(\frac{(\eta+1)\Delta t_M-t}{\Delta t_M}\right.$$

$$\left.+\frac{t-\eta\Delta t_M}{\Delta t_M}e^{-2A_M}\right)e^{-2\eta A_M}e^{-2\Delta\alpha\frac{t}{\Delta\tau}} \qquad (60)$$

(for spurious mode having no degeneration)

$$p(t)=\frac{\sigma_{K,d}{}^4|\Gamma-t|}{2l^2\Delta\tau^2}\left(\frac{(\eta+1)\Delta t_M-t}{\Delta t_M}\right.$$

---

* Larsen has already treated the influence of the TE_{01} reflection.⑴ This discussion is similar to his treatment, but the results are somewhat different.

$$+\frac{t-\eta\Delta t_M}{\Delta t_M}e^{-2A_M}\Big]e^{-2\eta A_M}e^{-2\Delta\alpha\frac{t}{\Delta\tau}} \qquad (61)$$

(for spurious mode having even-odd degeneration)
where

$$0 \leq \frac{t-\eta\Delta t_M}{\Delta t_M} \leq 1 \qquad (62)$$

$$\Delta t_M = \Delta\tau l_M \qquad (63)$$

$$\eta = 0, 1, 2, \cdots\cdots, N-1.$$

As it is complicated to treat the general case, we will discuss only the case of sufficiently large A_M. In this case it is sufficient to take only the first term of equation (60) or (61). It is assumed N is very large, so we can consider

$$|\Gamma| \gg |\Delta t_M| \quad \text{or} \quad L \gg l_M. \qquad (65)$$

Then equation (60) or (61) can be rewritten as:

$$p(t) = \frac{\sigma_K{}^4|\Gamma|}{l^2\Delta\tau^2}\,\frac{\Delta t_M-t}{\Delta t_M}e^{-2\Delta\alpha\frac{t}{\Delta\tau}} \qquad (66)$$

$$p(t) = \frac{\sigma_{K,d}{}^4|\Gamma|}{2l^2\Delta\tau^2}\,\frac{\Delta t_M-t}{\Delta t_M}e^{-2\Delta\alpha\frac{t}{\Delta\tau}} \qquad (67)$$

$$0 \leq \frac{t}{\Delta t_M} \leq 1. \qquad (68)$$

Integrating $p(t)$, the total power of the interference waves is given:

$$P = \frac{\sigma_K{}^4|\Gamma|}{l^2\Delta\tau^2}\,\frac{|\Delta t_M|}{2} = \frac{\sigma_K{}^4 N^2}{2(N/N_M)} \qquad (69)$$

(for spurious mode having no degeneration)

$$P = \frac{\sigma_{K,d}{}^4|\Gamma|}{2l^2\Delta\tau^2}\,\frac{|\Delta t_M|}{2} = \frac{\sigma_{K,d}{}^4\cdot N^2}{2(N/N_M)} \qquad (70)$$

(for spurious mode having even-odd degeneration)
where N_M is the number of waveguides between two adjoint mode filters, and we have assumed

$$2\Delta\alpha\frac{\Delta t_M}{\Delta t} = 2\Delta\alpha l_M = 2\Delta\alpha l N_M \ll 1 \qquad (71)$$

$$N_M = \frac{l_M}{l}. \qquad (72)$$

Comparing equation (69) or (70) with equation (53) or (55), we obtain a representation for the reduction factor of the interference wave power due to the inserting of mode filters:

$$H = \frac{N_H}{N}A \quad (A \text{ in Naper}). \qquad (73)$$

3. Echo Distortion in Frequency Modulation Systems

3.1. Derivation of the Echo Distortion Formula

Here we will apply the results of Section 2 to calculate the intermodulation noise which appears in the frequency division multiplex telephony channels when waveguides are used to transmit these channels with frequency modulation.

The intermodulation noise, which is caused by the interference waves in the waveguide transmission line, has the same property as the distortion caused by the echo (i.e. echo distortion). Therefore we will call it also "echo distortion."

The distortion formulas of a single echo are given at the point where the test tone level is 0 dBm as:(5)

$$D_2 = 8\pi^4 r_q{}^2 t^4\sigma^2 f^2\left(1-\frac{1}{2}\,\frac{f}{f_h}\right), \text{ [mW]} \qquad (74)$$

$$D_3 = 8\pi^6 r_p{}^2 t^6\sigma^4 f^2\left\{1-\frac{1}{3}\left(\frac{f}{f_h}\right)^2\right\}, \text{ [mW]} \qquad (75)$$

$$D_\infty = \gamma^2 K^2, \text{ (mW]} \qquad (76)$$

(for the highest frequency telephone channel)

where

D_2=approximate value of the second-order distortion for the short delayed echo.
D_3=approximate value of the third-order distortion for the short delayed echo.
D_∞=approximate value of the total distortion for the long delayed echo.
r_q=quadrature component of the echo.
r_p=in-phase compnent of the echo.
$r=\sqrt{\gamma_q^2+\gamma_p^2}$=amplitude ratio of the echo to the signal.
t=delay time of the echo.
σ=r.m.s. frequency deviation of the multiplex telephony signal.
ω=angular frequency of the carrier.
f=center frequency of a telephone channel.
f_h=center frequency of the highest frequency telephone channel.
K=function of σ/f_h given by Medhurst.(3)

As the mean delayed power spectrum of the interference waves in the multimode waveguide has a great many of spectral components, and these components may be considered to have random phase relation with the main wave, we will average equations (74) and (75) with respect to carrier phase, and will sum them up to get an expression for the total and average distortion due to a short delayed single echo.

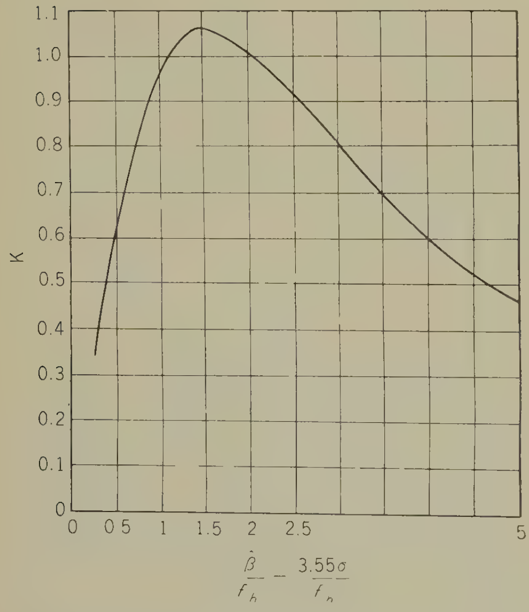


Fig. 2—Values of K given by Medhurst.

$$D_0=\overline{D}_2+\overline{D}_3=4\pi^4r^2t^4\sigma^2f^2\Big[\Big(1-\frac{1}{2}\frac{f}{f_h}\Big)+\pi^2t^2\sigma^2\Big\{1-\frac{1}{3}\Big(\frac{f}{f_h}\Big)^2\Big\}\Big]. \tag{77}$$

For simplicity, we will treat only the echo distortion which appears in the highest frequency telephone channel.

$$D_0=2\pi^4t^4\sigma^2f_h^2\Big(1+\frac{4}{3}\pi^2t^2\sigma^2\Big)\cdot r^2. \tag{78}$$

$$D_\infty=K^2\cdot r^2. \tag{79}$$

Though equations (78) and (79) are valid for very short and very long delayed echos respectively, we will also use them for medium delay time in order to simplify the treatment.

That is, we will define characteristic delay time t_c which satisfies

$$D_\infty=D_0 \tag{80}$$

and will use equation (78) for $t\geq t_c$, and equation (79) for $t\geq t_c$. The error caused by this approximation is considered to be smaller than about 3 dB.

To calculate the echo distortion due to a set of interference waves, we will replace r^2 with $p(t)$, which is defined in chapter 2, and integrate equation (78) or (79) with respect to t. We obtain

$$d=d_1+d_2 \tag{81}$$

$$d_1=\int_0^{t_c}D_0dt=\int_0^{t_c}2\pi^4t^4\sigma^2f_h^2\Big(1+\frac{4}{3}\pi t^2\sigma^2\Big)p(t)\,dt \tag{82}$$

$$d_2=\int_{t_c}^{\Gamma}D_\infty dt=\int_{t_c}^{\Gamma}K^2p(t)dt. \tag{83}$$

Where d is the total distortion power due to a mean delayed power spectrum of the interference waves, and d_1 and d_2 are distortion components due to interference waves for $t \leq t_c$ and $t \geq t_c$ respectively. The unit of d, d_1, and d_2 is milliwatts.

Equations (82) and (83) are rewritten as:

$$d_1 = \frac{\sigma_K{}^4 |\Gamma|}{l^2 \Delta\tau^2} \left\{ 2\pi^4 \sigma^2 f_h{}^2 \int_0^{t_c} t^4 e^{-2\left|\frac{\Delta\alpha}{\Delta\tau}\right| t} dt + \frac{8}{3} \pi^6 \sigma^4 f_h{}^2 \int_0^{t_c} t^6 e^{-2\left|\frac{\Delta\alpha}{\Delta\tau}\right| t} dt \right\} \tag{84}$$

$$d_2 = \int_{t_c}^{|\Gamma|} K^2 \frac{\sigma_K{}^4 (|\Gamma| - t)}{l^2 \Delta\tau^2} e^{-2\left|\frac{\Delta\alpha}{\Delta\tau}\right| t} dt \tag{85}$$

$$d_1 = \frac{\sigma_{K,d}{}^4 |\Gamma|}{2l^2 \Delta\tau^2} \left\{ 2\pi^4 \sigma^2 f_h{}^2 \int_0^{t_c} t^4 e^{-2\left|\frac{\Delta\alpha}{\Delta\tau}\right| t} dt + \frac{8}{3} \pi^6 \sigma^4 f_h{}^2 \int_0^{t_c} t^6 e^{-2\left|\frac{\Delta\alpha}{\Delta\tau}\right| t} dt \right\} \tag{86}$$

$$d_2 = \int_{t_c}^{|\Gamma|} K^2 \frac{\sigma_{K,d}{}^4 (|\Gamma| - t)}{2l^2 \Delta\tau^2} e^{-2\left|\frac{\Delta\alpha}{\Delta\tau}\right| t} dt \tag{87}$$

where equations (84) and (85) are for spurious modes having no degeneration, and equations (86) and (87) are for spurious modes having even-odd degeneration. In equations (84) and (86) we have used the approximation:

$$\Gamma - t \doteqdot \Gamma \tag{88}$$

because $t \leq t_c \ll |\Gamma|$.

We will define a function $E_n(a)$ as follows:

$$E_n(a) = (n+1) \int_0^1 u^n e^{-au} du. \tag{89}$$

Where a is a non-negative number and n is a positive integer. Characteristics of $E_n(a)$ are examined in Appendix 6.3.

The factor in braces of equation (84) and (86) is written as R. Using equation (89), we obtain

$$R = Q_1 E_4\left(2\left|\frac{\Delta\alpha}{\Delta\tau}\right| t_c\right) + Q_2 E_6\left(2\left|\frac{\Delta\alpha}{\Delta\tau}\right| t_c\right), \tag{90}$$

where

$$Q_1 = \frac{2}{5} \pi^4 \sigma^2 f_h{}^2 t_c{}^5 \tag{91}$$

$$Q_2 = \frac{8}{21} \pi^6 \sigma^4 f_h{}^2 t_c{}^7. \tag{92}$$

Equations (85) and (87) can be calculated assuming:

$$2|\Gamma| \cdot \left|\frac{\Delta\alpha}{\Delta\tau}\right| = 2A \gg 1 \tag{93}$$

$$|\Gamma| \gg t_c \tag{94}$$

$$e^{-2\left|\frac{\Delta\alpha}{\Delta\tau}\right| |\Gamma|} \ll e^{-2\left|\frac{\Delta\alpha}{\Delta\tau}\right| t_c}. \tag{95}$$

Then we obtain

$$\int_{t_c}^{|\Gamma|} (|\Gamma| - t) e^{-2\left|\frac{\Delta\alpha}{\Delta\tau}\right| t} dt = \frac{1}{2} \frac{\Gamma^2}{A} e^{-2\left|\frac{\Delta\alpha}{\Delta\tau}\right| t_c}. \tag{96}$$

Summarizing the results derived here, we obtain

$$d_1 = \frac{\sigma_K{}^4 \cdot N}{l |\Delta\tau|} R \tag{97}$$

$$d_2 = \frac{K^2 \sigma_K{}^4 N}{2l \Delta\alpha} e^{-2\left|\frac{\Delta\alpha}{\Delta\tau}\right| t_c} \tag{98}$$

for spurious modes having no degeneration, and

$$d_1 = \frac{\sigma_{K,d}{}^4 \cdot N}{2l |\Delta\tau|} R \tag{99}$$

$$d_2 = \frac{K^2 \cdot \sigma_{K,d}{}^4 \cdot N}{4l \Delta\alpha} e^{-2\left|\frac{\Delta\alpha}{\Delta\tau}\right| t_c} \tag{100}$$

for spurious modes having even-odd degeneration.

Using the same method as in the mode conversion and reconversion, the formulas for the signal mode reflections are calculated as:

$$d_1 = \frac{\sigma_R{}^4 \cdot N}{2l\tau} R \tag{101}$$

$$d_2 = \frac{K^2 \sigma_R{}^4 N}{4l\alpha} e^{-2\frac{\alpha}{\tau} t_c} \tag{102}$$

where R is defined as:

$$R = Q_1E_4\left(2\frac{\alpha}{\tau}t_c\right)+Q_2E_6\left(2\frac{\alpha}{\tau}t_c\right) \qquad (103)$$

as in equation (90).

3.2. Effects of Spurious Mode Loss and Frequency Deviation on the Distortion Level

Before the treatments of these problems, we will first examine how much frequency deviation is required for practical communication.

We will take an example of noise allotment which satisfies the C.C.I.R. recommendation of 1954*, and will calculate the minimum frequency deviation required for satisfying it.

- Total noise −57 dBm (per 280 km)
  - thermal noise, −60 dBm (per 280 km)
  - interference noise, −70 dBm (per 280 km)
  - intermodulation noise, −60 dBm (per 280 km)
    - nonlinear distortion −71.4 dBm (per 280 km)
    - delay distortion −63 dBm (per 280 km)
    - echo distortion −63 dBm (per 280 km)

Where these values indicate the allowable noise power at the point where the test tone level is 0 dBm.

The thermal noise power appearing in a telephone channel through transmission of a repeating span is given as$^{(5)}$:

$$D_{Th} = \frac{(KTb)}{\left(\frac{P_R}{F_N}\right)}\left(\frac{f}{s_0}\right)^2 \qquad (104)$$

where

K = Boltzmann constant, 1.37×10^{-23} joule/°K

T = absolute temperature of system ($\doteqdot 239$ °K)

b = bandwidth of each telephone channel = 3.1 kc

$KTb = 1.24\times10^{-14}$ mW = −139.1 dBm

F_N = noise figure of repeater

P_R = received R.F. power

f = center frequency of telephone channel

s_0 = r.m.s. frequency deviation corresponding to the test tone level

For simplicity, we will consider only the highest frequency telephone channel in a frequency modulation system without emphasis.

Rewriting equation (104), we obtain:

$$\left(\frac{s_0}{f_h}\right) = \frac{(KTb)}{\left(\frac{P_R}{F_N}\right)D_{Th}}. \qquad (105)$$

Here we will assume:*

$$\frac{P_R}{F_N} = -50\text{dBm} \qquad (106)$$

for a repeating span of 20 km. As the allowable thermal noise is

$$D_{Th} = -71.46 \text{ dBm}$$

for a repeating span of 20 km, we obtain

$$\frac{s_0}{f_h} = 0.132 \quad \text{or} \quad s_0 = 0.132 f_h. \qquad (107)$$

Or, if we use the number of telephone channels N_c,

$$s_0 = 0.528\times10^{-3}\,N_c, \quad (s_0 \text{ in Mc}) \qquad (108)$$

where

---

* Total transmission length: 2500 km,
Transmission length between modulator and demodulator: 280 km,
Allowable average noise power per telephone channel: 7500 pW.

* This assumption corresponds, for example, to a transmitting power of 20 dBm, transmission loss of 45 dB for a 20 km repeating span, and a noise figure of 25 dB. Therefore, it is considered that P_R/F_N of −50 dBm is a fairly conservative value.

$$f_h = 4\times10^{-3}N_c, \quad (f_h \text{ in Mc}) \qquad (109)$$

σ, which is the r.m.s. frequency deviation of multiplex telephone signal, is related to s_0 and N_c by⁽⁵⁾:

$$\sigma = s_0\times10^{\frac{-15+\log_{10}N_c}{20}} = \frac{s_0\sqrt{N_c}}{5.62}. \qquad (110)$$

Then, using equation (108), we obtain:

$$\sigma = 9.41\times10^{-5}N_c^{\frac{3}{2}}. \quad (\sigma \text{ in Mc}) \qquad (111)$$

Or, if we use $\hat{\beta}$, which is the peak deviation of multiplex telephone signal, we obtain:

$$\hat{\beta} = 3.55\sigma = 3.34\times10^{-4}N_c^{\frac{3}{2}} \qquad (112)$$

or,

$$\frac{\hat{\beta}}{f_h} = 0.0834\sqrt{N_c} = 1.32\sqrt{f_h}. \qquad (113)$$

As the r.m.s. or peak deviation calculated here is considered to be the minimum deviation required for satisfying the thermal noise requirement of equation (104), in the present paper we will call them the minimum r.m.s. or minimum peak deviation of the system.

Furthermore, in order to get a rough estimation of band requirement, we will assume that the bandwidth required for transmitting a frequency modulated ratio frequency signal is given by:

$$B = 3(2\hat{\beta}+2f_h) \qquad (114)$$

including the guard band.

Now, we will consider the characteristics of the parameters appearing in the equation of echo distortion.

K is a function of $\hat{\beta}/f_h$ given, in Fig. 2, by Medhurst for $\hat{\beta}/f_h\leq5$. For the large values of $\hat{\beta}/f_h$, we can use Albersheim's formula.

$$K=5.8\left(\frac{f_h}{\hat{\beta}}\right)^{1.5}\exp\left[-2.88\left(\frac{f_h}{\hat{\beta}}\right)^{2}\right],$$

$$\text{for } \frac{\hat{\beta}}{f_h}\geq5. \qquad (115)$$

Furthermore, for $\hat{\beta}/f_h\geq10$, equation (105) is approximated by:

$$K = 5.8\left(\frac{f_h}{\hat{\beta}}\right)^{1.5} \qquad (116)$$

and t_c is the characteristic delay time which satisfies

$$K^2 = 2\pi^4t_c^4\sigma^2f_h^2\left(1+\frac{4}{3}\pi^2t_c^2\sigma^2\right) \qquad (117)$$

from equations (78), (79), and (80).

Rewriting equation (117), we obtain

$$\left(\frac{\sigma}{f_h}\right)K = \sqrt{2}\,(\pi t_c\sigma)^2\sqrt{1+\frac{4}{3}(\pi t_c\sigma)^2}\,. \qquad (118)$$

As the left-hand side of equation (118) is function of only σ/f_h or $\hat{\beta}/f_h$, $t_c\sigma$ is considered to be function of only σ/f_h or $\hat{\beta}/f_h$. This relation is calculated numerically and plotted in Fig. 3.

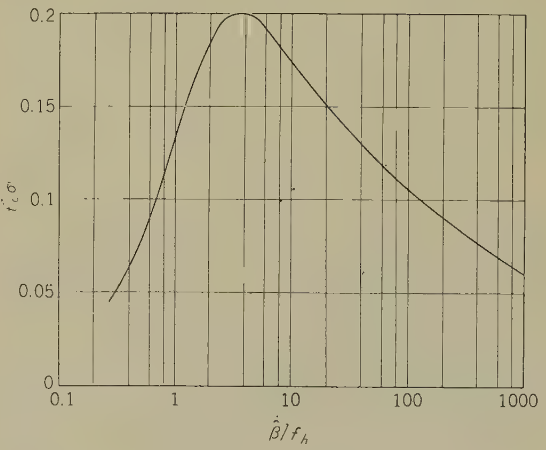


Fig. 3—$t_c\sigma\sim\hat{\beta}/f_h$.

The relation of Fig. 3 can be rewritten to the relation between t_cf_h and $\hat{\beta}/f_h$ because

$$t_c f_h = (t_c\sigma)\frac{f_h}{\sigma} = (t_c\sigma)3.55\frac{f_h}{\hat{\beta}}.$$

Fig. 4 shows this relation.

When

$$\pi t_c\sigma \ll 1, \qquad \frac{\hat{\beta}}{f_h} \gg 1 \tag{119}$$

equation (118) can be approximated by

$$t_c\sigma = 0.342\left(\frac{\hat{\beta}}{f_h}\right)^{-\frac{1}{4}} \tag{120}$$

or

$$t_c f_h = 1.215\left(\frac{\hat{\beta}}{f_h}\right)^{-1.25}. \tag{121}$$

We will consider Q_1 and Q_2. From equations (91) and (92) we can write:

$$f_h Q_1 = 5.70\left(\frac{f_h}{\hat{\beta}}\right)^3(\pi t_c\sigma)^5 \tag{122}$$

$$f_h Q_2 = 5.43\left(\frac{f_h}{\hat{\beta}}\right)^3(\pi t_c\sigma)^7. \tag{123}$$

Therefore, $f_h Q_1$ and $f_h Q_2$ are functions of only $\hat{\beta}/f_h$. Their values are calculated in Fig. 5. For sufficiently large $\hat{\beta}/f_h$, we get:

$$f_h Q_1 = 8.16\left(\frac{\hat{\beta}}{f_h}\right)^{-4.25} \tag{124}$$

$$f_h Q_2 = 8.98\left(\frac{\hat{\beta}}{f_h}\right)^{-4.75}. \tag{125}$$

Now we will consider the characteristics of the echo distortion, taking advantage of the formulas calculated up to this point.

For copper-pipe waveguides, $\Delta\alpha$ is so small for the main unwanted modes that we can consider roughly:

$$E_4\left(2\left|\frac{\Delta\alpha}{\Delta\tau}\right|t_c\right) \doteqdot 1$$

$$E_6\left(2\left|\frac{\Delta\alpha}{\Delta\tau}\right|t_c\right) \doteqdot 1$$

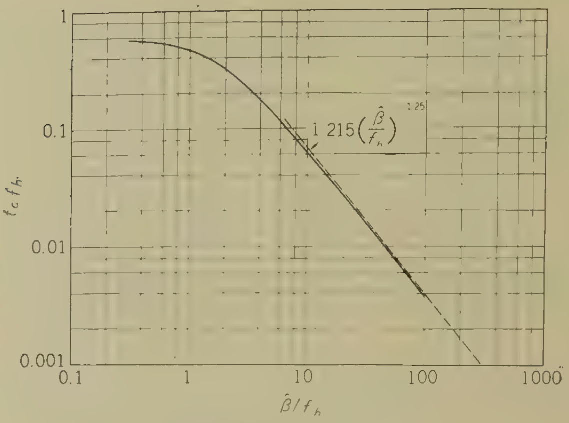


Fig. 4—$t_c f_h \sim \hat{\beta}/f_h$.

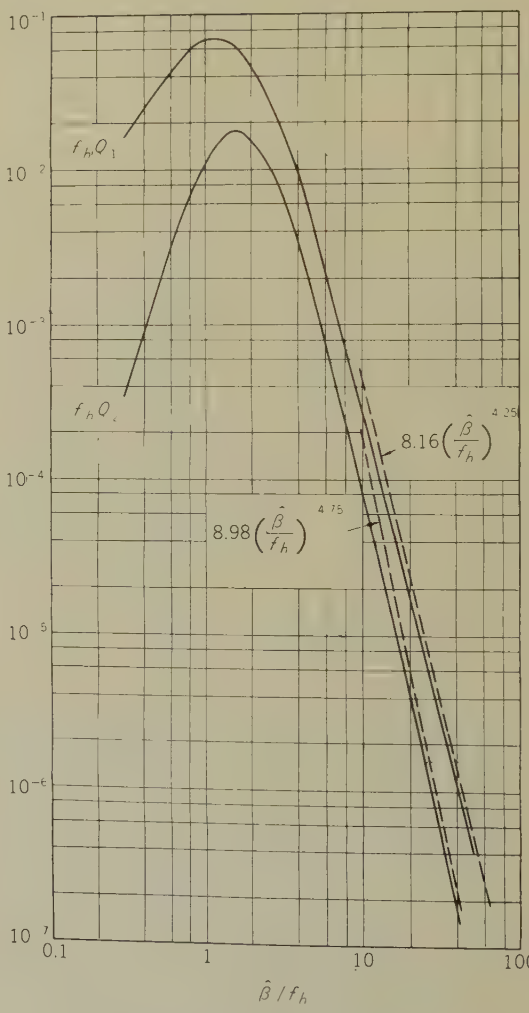


Fig. 5—$f_h Q_1$ and $f_h Q_2 \sim \hat{\beta}/f_h$.

$$e^{-2}\left|\frac{\Delta\alpha}{\Delta\tau}\right| t_c \doteqdot 1.$$

Furthermore, in the case of small $\Delta\alpha$, it can be observed that d_1 is smaller than d_2 in practical circular waveguides. Then, neglecting d_1, we obtain a rough approximation for the total distortion d:

$$d \doteqdot \frac{K^2 \cdot \sigma_K{}^4 \cdot N}{2l\Delta\alpha} \quad \cdots\cdots \text{for spurious modes having no degeration} \tag{126}$$

$$d \doteqdot \frac{K^2 \cdot \sigma_{K,d}{}^4 \cdot N}{4l\Delta\alpha} \quad \cdots\cdots \text{for spurious modes having even-odd degeneration.} \tag{127}$$

For large $\hat{\beta}/f_h$, equation (126) or (127) is proportional to $(\hat{\beta}/f_h)^{-3}$, because K is proportional to $(\hat{\beta}/f_h)^{-1.5}$.

For waveguides of large $\Delta\alpha$, for example for the case of helix waveguides, we can assume $2|\Delta\alpha/\Delta\tau|t_c$ is sufficiently large. Then we obtain

$$E_4\left(2\left|\frac{\Delta\alpha}{\Delta\tau}\right| t_c\right) \doteqdot \frac{5!}{2^5}\left(\left|\frac{\Delta\alpha}{\Delta\tau}\right| t_c\right)^{-5} \tag{128}$$

$$E_6\left(2\left|\frac{\Delta\alpha}{\Delta\tau}\right| t_c\right) \doteqdot \frac{7!}{2^7}\left(\left|\frac{\Delta\alpha}{\Delta\tau}\right| t_c\right)^{-7} \tag{129}$$

from Appendix 6.3.

Therefore we can consider

$$d_1 \gg d_2$$

$$Q_1E_4\left(2\left|\frac{\Delta\alpha}{\Delta\tau}\right| t_c\right) \gg Q_2E_6\left(2\left|\frac{\Delta\alpha}{\Delta\tau}\right| t_c\right).$$

And the total distortion d is given by:

$$d \doteqdot \frac{\sigma_K{}^4 \cdot N}{l|\Delta\tau|}\hat{R} \quad \cdots\cdots \text{for spurious modes having no degeneration} \tag{130}$$

$$d \doteqdot \frac{\sigma_{K,d}{}^4 \cdot N}{2l|\Delta\tau|}\hat{R} \quad \cdots\cdots \text{for spurious modes having even-odd degeneration} \tag{131}$$

where $\hat{R}$ is given by:

$$\hat{R} \doteqdot 11.6\hat{\beta}^2 f_h{}^2\left|\frac{\Delta\alpha}{\Delta\tau}\right|^{-5}. \tag{132}$$

It must be noticed that the effects of attenuation difference $\Delta\alpha$ and frequency deviation $\hat{\beta}$ on the echo distortion are different according as $\Delta\alpha$ is very large or very small. When $\Delta\alpha$ is very small, d is proportional to $\Delta\alpha^{-1}$ and $\hat{\beta}^{-3}$. On the other hand, when $\Delta\alpha$ is very large, d is proportional to $\Delta\alpha^{-5}$ and $\hat{\beta}_2$.

In helix waveguides echo distortion caused by spurious modes other than TE_{0n} is much reduced because $\Delta\alpha$ is very large for these modes. We will examine how much $\Delta\alpha$ is required for practical helix waveguides. In Fig. 6, equation (90) is calculated as a function of $|\Delta\alpha/\Delta\tau|$. In the calculation we have assumed that $f_h = 8$ or 16 Mc, and peak deviation takes minimum value given by equation (112). Waveguide diameter is 51 mm, and carrier frequency is 50 Gc. Arrows in Fig. 6 indicate $|\Delta\alpha/\Delta\tau|$ is for spurious modes in the copper-pipe waveguide. The spurious modes, which cause the greater portion of the distortion in the coppar-pipe waveguide and can be eliminated in the helix waveguide, are thought to be TE_{12} and TE_{22} modes. Because $|\Delta\alpha/\Delta\tau|$ is about 10^{-2} Np/mμs for these modes, we arrive at the following conclusions from Fig. 6.

In helix waveguide having $\Delta\alpha$ 100 times as large as copper-pipe waveguides is used the echo distortion may be made much smaller.

Or, if we define a new parameter ξ for waveguides other than copper-pipe waveguides as:

$$\xi = \frac{[\Delta\alpha]}{[\Delta\alpha]_{\text{copper}}} \tag{133}$$

it is said that ξ must be as large as 10^2 in order to make the echo distortion sufficiently small.

Here we will calculate some examples numerically, and examine to what extent the approximate formulas (126), (129), and (131) are valid in the case of practical circular waveguides.

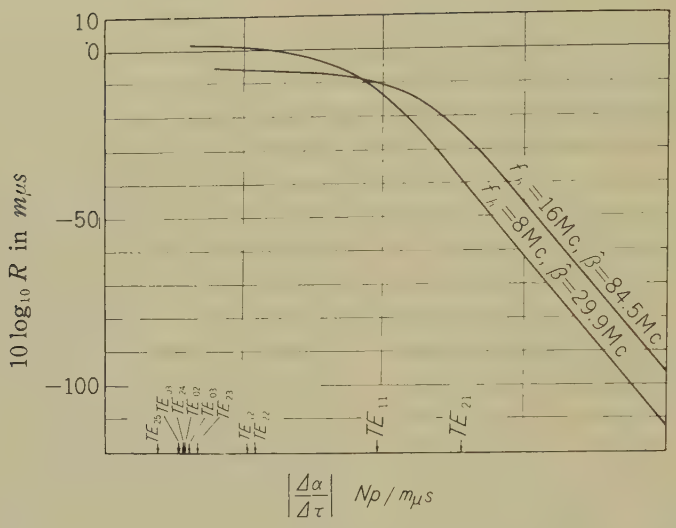


Fig. 6—Effect of $|\Delta\alpha/\Delta\tau|$ on R, $f_0=50$ Gc, $2a=51$ mm. Arrows indicate $|\Delta\alpha/\Delta\tau|$ ofmodes in copper-pipe circulate waveguide. Where the unit of R is $m\mu s$.

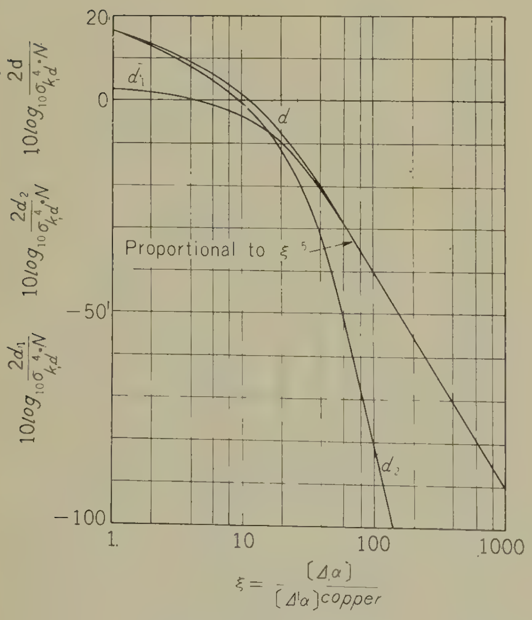


Fig. 7—d_1, d_2, and d due to the TE_{11} mode, as a function of ξ. $f_0=50$ Gc, $l=5$ m, $f_h=16$ Mc, $\hat{\beta}=84.5$ Mc.

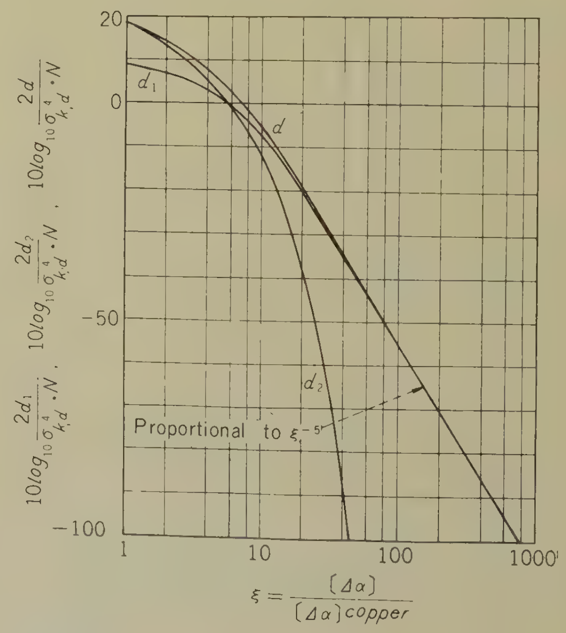


Fig. 8—d_1, d_2, and d due to the TE_{12} mode, as a function of ξ. $f_0=50$ Gc, $l=5$ m, $f_h=8$ Mc, $\hat{\beta}=29.9$ Mc.

Figs. 7 and 8 indicate the echo distortion due to the TE_{12} mode where the abscissa indicates ξ, and frequency deviation takes the minimum value as indicated in equation (111) or (112). In these examples, while d_2 is much larger than d_1 where ξ is close to 1, d_1 is much larger than d_2 for large ξ, and d is proportional to ξ^{-5} or $[\Delta\alpha]^{-5}$ for $\xi \geq 10$.

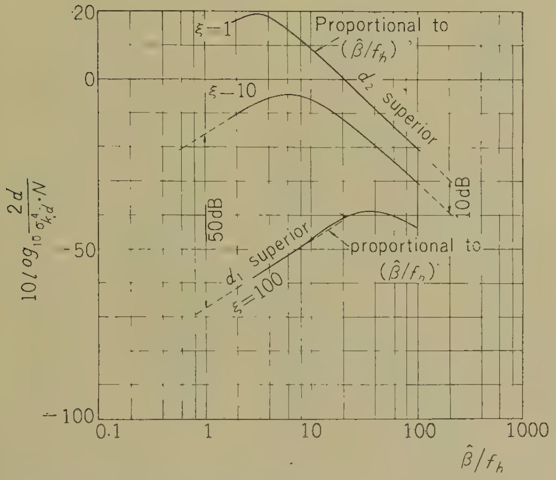


Fig. 9—Effect of frequency deviation on echo distortion. $f_0=50$ Gc, $l=5$ m, $f_h=8$ Mc, $2a=51$ mm. Distortion due to the TE_{12} mode is illustrated.

Fig. 9 shows the effect of the frequency deviation. The echo distortion due to the TE_{12} mode is calculated with respect to $\hat{\beta}/f_h$ from 1 to 100 and for three values of ξ: 1, 10, and 100.

When ξ is 1, d_2 is much larger than d_1, where $\hat{\beta}/f_h$ exceeds 10, and equation (127) is utilized. Therefore, distortion power is proportional to $(\hat{\beta}/f_h)^{-3}$ and $(\Delta\alpha)^{-1}$.

When ξ is 100, d_1 is much larger than d_2, where $\hat{\beta}/f_h$ does not exceed 10, and equation (131) is utilized. Therefore, distortion power is proportional to $(\hat{\beta}/f_h)^2$ and $(\Delta\alpha)^{-5}$.

Generally, echo distortion due to a spurious mode varies with the peak deviation $\hat{\beta}$ in the following manner: the curve of echo distortion vs. $\hat{\beta}$ is mountain shaped. On the right slope, distortion due to long delayed interference waves dominant and almost proportional to $(\Delta\alpha)^{-1}$ and $\hat{\beta}^{-3}$. On the left slope, distortion due to short delayed interference waves is dominant and almost proportional to $(\Delta\alpha)^{-5}$ and $\hat{\beta}^2$.

3.3. Allowable Tolerances for Copper-Pipe Waveguides

In circular waveguides which are used for long distance communication, there are many spurious modes.

At the present time the spurious modes which are considered to be harmful to signal transmission are:

TE_{0n} group, TE_{1n} group, TE_{2n} group

where the TM_{11} mode is excluded because it has the same phase velocity as the TE_{01} mode and cannot produce interference waves.

TE_{0n} modes are excited by the TE_{01} mode through waveguide junctions having a diameter difference.

TE_{1n} modes are excited by the TE_{01} mode through waveguide junctions having either tilt, offset, or both.

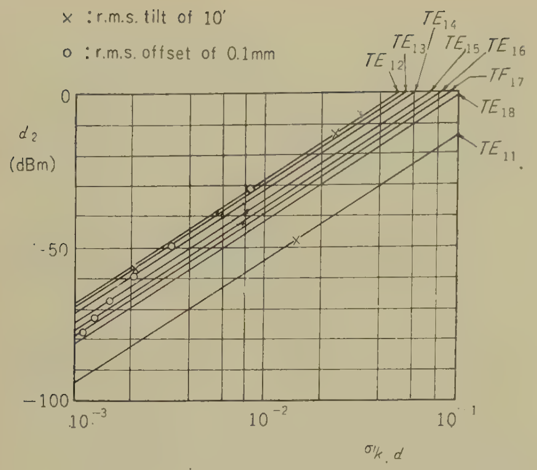
TE_{2n} modes are excited by the TE_{01} mode through waveguide junctions having elliptical deformation,

Mode coupling coefficients due to these deformations are shown in Appendix 6.1.

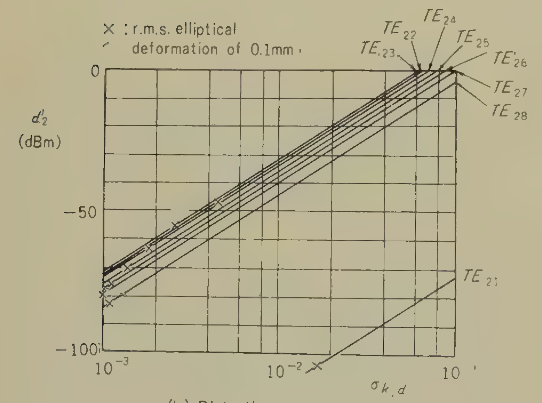
We will calculate the echo distortion due to each spurious mode as a function σ_K or $\sigma_{K,d}$. Here we will assume:

$f_0=50$ Gc ....carrier frequency
$2a=51$ mm ....waveguide diameter
$l=5$ m......length of unit waveguide
$L=20$ km ....total length of transmission line
$N_c=2000$ or 4000....number of telephone channels
$f_h=8$ Mc or 16 Mc ..bandwidth of multiplex telephone signal.

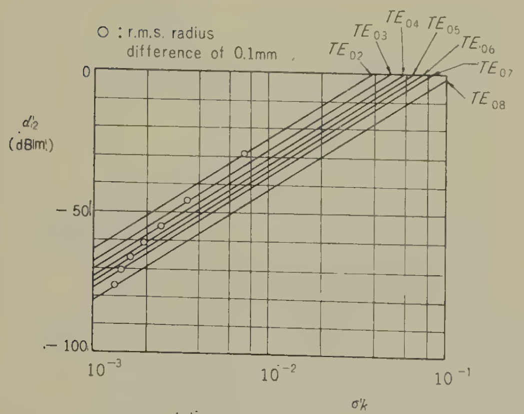
In Fig. 10 and 11, echo distortion is illustrated as a function of σ_K or $\sigma_{K,d}$, where peak deviation is set to the minimum value indicated in equation (112). To show a practical



(a) Distortion due to TE $_{2n}$

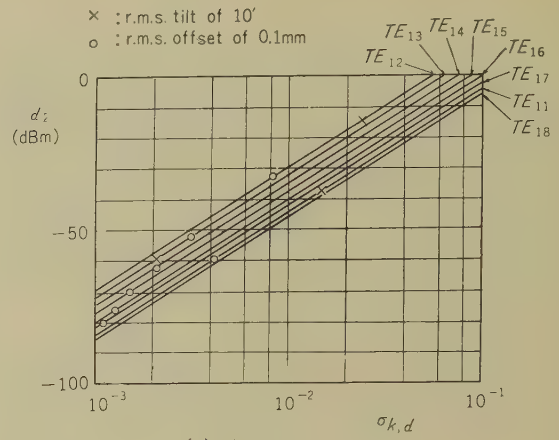


(b) Distortion due to TE $_{2n}$

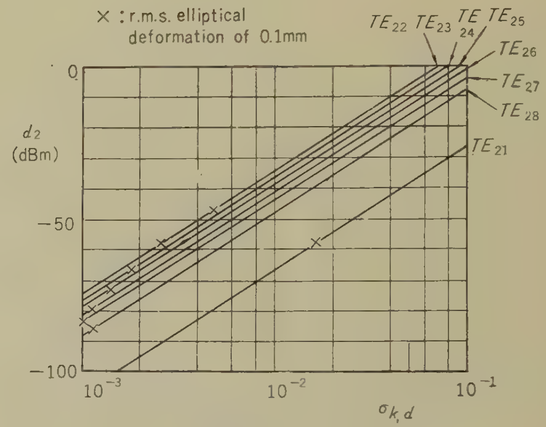


(c) Distortion due to TE $_{0n}$

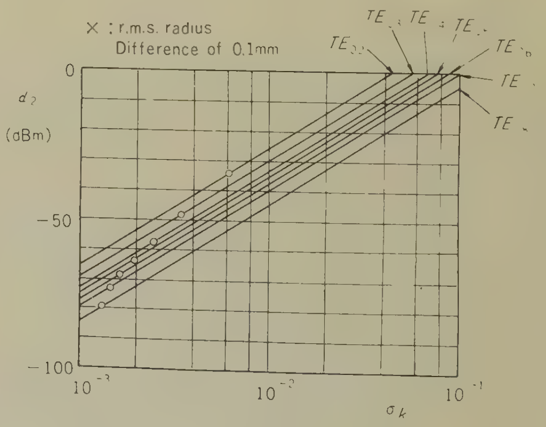
Fig. 10—Echo distortion appearing in top telephone channel due to spurious modes in copper-pipe waveguides. $N_c=2000$, $f_h=8$ Mc, $\hat{\beta}=29.9$ Mc, $l=5$ m, $2a=51$ mm, $L=20$ km.



(a) Distortion due to TE $_{1n}$



(b) Distortion due to TE $_{2n}$



(c) Distortion due to TE $_{0n}$

Fig. 11—Echo distortion appearing in top telephone channel due to spurious modes in copper-pipe waveguides. $N_c=4000$, $f_h=16$ Mc, $\hat{\beta}=84.5$ Mc, $l=5$ m, $2a=51$ mm, $L=20$ km.

example, echo distortion due to mode conversion at junctions is indicated by ○'s and ×'s where

standard deviation of junction tilt $=10$ minutes
standard deviation of junction offset $=0.1$ mm
standard deviation of radius difference at junction $=0.1$ mm
standard deviation of elliptical deformation junction $=0.1$ mm.

Because we have not sufficient data to treat the case with continuous mode conversion, we will discuss only the case with discrete mode conversion at the junctions. Then it is observed that the TE_{02}, TE_{12}, or TE_{22} modes produce the larger portion of the distortion in the TE_{0n}, TE_{1n}, and TE_{2n} mode groups. The other modes produce only a small portion of the distortion, because the TE_{11} and TE_{21} modes have very large $2|\Delta\alpha/\Delta\tau|t_c$, and the TE_{03}, TE_{04}, $\cdots$; TE_{13}, TE_{14}, $\cdots$, TE_{23}, TE_{24}, $\cdots$ modes have very small conversion coefficients.

The relation between d_2 and σ_K or $\sigma_{K,d}$ for the main three modes is as follows:

For $N_c=2000$
TE_{12}: $10 \log_{10} d_2 = 40 \log_{10} \sigma_{K,d}+51.6$
TE_{22}: $10 \log_{10} d_2 = 40 \log_{10} \sigma_{K,d}+47.2$
TE_{02}: $10 \log_{10} d_2 = 40 \log_{10} \sigma_{K}+56.4$
For $N_c=4000$
TE_{12}: $10 \log_{10} d_2 = 40 \log_{10} \sigma_{K,d}+49.8$
TE_{22}: $10 \log_{10} d_2 = 40 \log_{10} \sigma_{K,d}+45.7$
TE_{02}: $10 \log_{10} d_2 = 40 \log_{10} \sigma_{K}+53.9$.

Allowable echo distortion per repeating span is -74.5 dBm, if we use the example of noise allotment shown in section 3.2. We will distribute -74.5 dBm to the three main modes as follows:

TE_{12}: -77.5 dBm
TE_{22}: -80.5 dBm
TE_{02}: -80.5 dBm.

We have alloted 3 dB more power to the TE_{12} mode than to the other modes, because the interference waves of the TE_{12} mode are produced by two mechanisms—offset and tilt. Then we obtain Table 1, which illustrates the allowable standard deviation of mode coupling coefficients when copper-pipe waveguides are used for transmitting multiplex telephone signals with frequency modulation.

Table 1

ALLOWABLE STANDARD DEVIATION OF MODE COUPLING COEFFICIENTS AT WAVEGUIDE JUNCTION WHEN COPPER-PIPE WAVEGUIDES ARE USED FOR TRANSMITTING MULTIPLEX TELEPHONE SIGNALS WITH FREQUENCY MODULATION

$2a=51$ mm, $l=5$ m, $L=20$ km, $f_0=50$ Gc
$\hat{\beta}$ is selected to be minimum frequency deviation

	$N_c=2000$, $f_h=8$ Mc	$N_c=2000$, $f_h=16$ Mc
TE_{12}	$\sigma_{K,d}=5.89\times10^{-4}$	$\sigma_{K,d}=6.60\times10^{-4}$
TE_{22}	$\sigma_{K,d}=6.45\times10^{-4}$	$\sigma_{K,d}=6.91\times10^{-4}$
TE_{02}	$\sigma_{K}=3.80\times10^{-4}$	$\sigma_{K}=4.36\times10^{-4}$

We have larger allowances in the case of 4000 channels than in the case of 2000 channels. The reason is that $\hat{\beta}/f_h$ is larger in the former case than in the later case.

When we calculate waveguide torelances using Table 1, we must take care of the case of the TE_{12} mode. Mode conversion between the TE_{01} and TE_{12} modes is caused by two mechanisms—offset and tilt. Because these two mecanisms can be considered to be independent of each other, we obtain:

$$\sigma_{K,d}^2(\text{tilt})+\sigma_{K,d}^2(\text{offset}) = \sigma_{K,d}^2(\text{total}). \quad (134)$$

If we allot equal tolerance to the offset and tilt, we obtain:

$$\sigma_{K,d}(\text{tilt}) = \sigma_{K,d}(\text{offset}) = 4.17\times10^{-4}$$

$$\text{for} \quad N_c = 2000$$

$$\sigma_{K,d}(\text{tilt}) = \sigma_{K,d}(\text{offset}) = 4.67\times10^{-4}$$

$$\text{for} \quad N_c = 4000.$$

By referring to Appendix 6.1., we can obtain Table 2 for allowable tolerances of waveguide junctions.

The tolerances in Table 2 are very small. We can conclude that it is very difficult to

Table 2

ALLOWABLE STANDARD DEVIATION OF WAVEGUIDE TOLERANCES AT JUNCTION WHEN COPPER-PIPE WAVEGUIDES ARE USED FOR TRANSMITTING MULTIPLEX TELEPHONE SIGNALS WITH FREQUENCY MODULATION

$2a=51$ mm, $l=5$ m, $L=20$ km, $f_0=50$ Gc
$\hat{\beta}$ is selected to be minimum frequency deviation

	$N_c=2000$	$N_c=4000$
Standard deviation of junction tilt	0.17′	0.20′
Standard deviation of junction offset	5.0 μ	5.6 μ
Standard deviation of junction elliptical deformation	13 μ	14 μ
Standard deviation of radius difference at junction	6.2 μ	7.1 μ

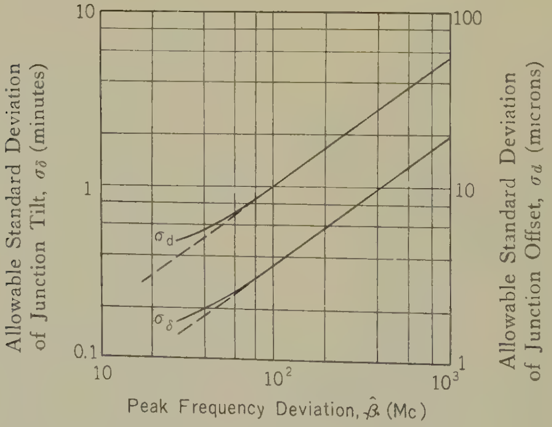
use copper-pipe waveguides for frequency modulation communication with minimum peak deviation.

We will now discuss the case of larger peak deviation. Waveguide tolerance increases with $\hat{\beta}^{5/4}$ in copper-pipe waveguides, because echo distortion decreases with $\hat{\beta}^{-3}$ for large $\hat{\beta}$ and increases with $\sigma_K{}^4$ or $\sigma_{K,d}{}^4$.

Calculating the tolerances of tilt and offset for $N_c=2000$, we obtain Fig. 12 (a).

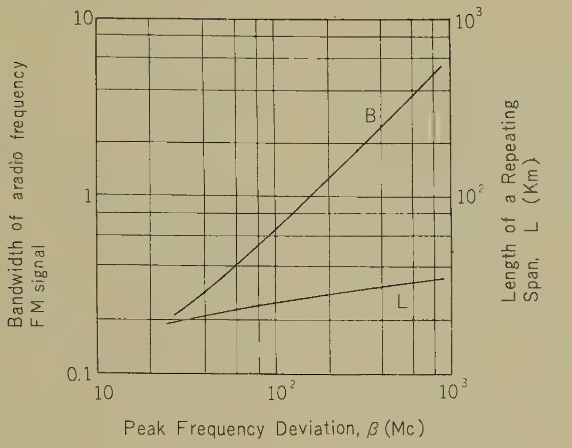
In Fig. 12 (b), the length of a repeating span and bandwidth required for transmitting a frequency modulated radio frequency signal are illustrated. The length of a repeating span is increased with the increase of peak frequency deviation because of equation (104).

From Fig. 12, we may obtain some information to compare frequency modulation with pulse code modulation. A pulse code modulation system suggested by S. E. Miller(6) has a bandwidth of 500 Mc and bit rates of 160 Mc. It permits transmission of 10 Mc of baseband information, i.e., two television signals of commercial quality or approximately 2000 telephone channels. If we take a bandwidth of 500 Mc in Fig. 12, the peak frequency deviation and the length of a repeating span are observed to be 75 Mc and



(a) Allowable waveguide tolerances for large frequency deviation.

Fig. 12—Frequency modulation transmission of multiplex telephone signal through copper-pipe waveguides. Case of large frequency deviation. $f_h=8$ Mc, $2a=51$ mm, $l=5$ m, $f_0=50$ Gc. Signal mode loss is assumed to be 2 dB/km.



(b) Bandwidth and repeating span for large frequency deviation.

Fig. 12—Frequency modulation transmission of multiplex telephone signal through copper-pipe waveguides. Case of large frequency deviation. $f_h=8$ Mc, $2a=51$ mm, $l=5$ m, $f_0=50$ Gc. Signal mode loss is assumed to be 2 dB/km.

24 km respectively. Furthermore, standard deviation of tilt and offset at each waveguide junction must not be larger than 0.2 minute and 8.5 microns respectively. We have little knowledge about the waveguide tolerances in the case of pulse code modulation, but they are thought to be much larger than in the case of frequency modulation.

3.4. Allowable Tolerances for Helix Waveguides

We will assume that the influence of the spurious modes other than the TE_{0n} group of modes is negligibly small compared to the echo distortion in helix waveguides, because $\Delta\alpha$ can be made sufficiently large for these spurious modes.

Figs. 10 (c) and 11 (c) are also applicable to helix waveguides under the same conditions as in section 3.4.

We will attribute all echo distortion power to the TE_{02} mode because the contributions of the TE_{03}, $TE_{04}\cdots$modes are very small.

Furthermore, we will use the standard deviation of waveguide diameter σ_D in place of the standard deviation of radius difference of waveguides at junction σ_δ to show the waveguide tolerance. The relation between σ_D and σ_δ is:

$$\sigma_\delta=\frac{1}{\sqrt{2}}\sigma_D. \tag{135}$$

Therefore we can obtain σ_K from the following equations:

$$10\log_{10} d_2=40\log_{10}\sigma_K+56.4 \quad \text{for} \quad N_c=2000$$

$$10\log_{10} d_2=40\log_{10}\sigma_K+53.9 \quad \text{for} \quad N_c=4000$$

$$10\log_{10} d_2=-74.5.$$

σ_K is rewritten as σ_δ or σ_D referring to Appendix 6.1 and equation (135). The results are shown in Table 3.

The results in Table 3 are for the case of minimum peak deviation given in equation (112). For the larger peak deviations we get the larger waveguide tolerances.

In figure 13, σ_D is calculated as a function

Table 3

ALLOWABLE TOLERANCES OF HELIC WAVEGUIDES WHICH ARE USED FOR FREQUENCY MODULATION TRANSMISSION OF MULTIPLEX TELEPHONE SIGNALS

$2a=51$ mm, $l=5$ m, $L=20$ km, $f_0=50$ Gc

$\hat{\beta}$ is selected to be minimum frequency deviation.

	$N_c=2000$	$N_c=4000$
Allowable standard deviation of coupling coefficient at junction	5.37×10^{-4}	6.17×10^{-4}
Allowable standard deviation of waveguide diameter	$12.5\ \mu$	$14.3\ \mu$

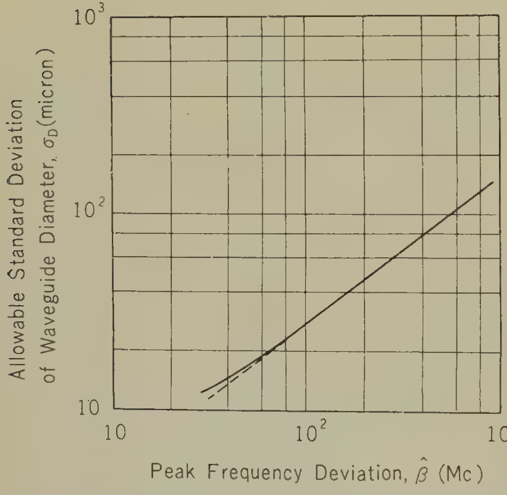


Fig. 13—Allowable tolerance of helix waveguidss for large peak frequency deviation.

of $\hat{\beta}$ for $N_c=2000$. Figure 12 (b) is also applicable to the case shown in figure 13.

The allowable standard deviation of waveguide diameter becomes about 22 microns, when we take $\hat{\beta}$ of 75 Mc, which is also taken for the example in section 3.3 which compares frequency modulation with the pulse code modulation.

3.5. Suppression of the Echo Distortion with Mode Filters

We have discussed the characteristics of interference waves in waveguide lines with mode filters, and calculated the mean delayed power spectrum of these interference waves in section 2.5.

Here, we will discuss the suppression of echo distortion with mode filters when the waveguide line is used for the frequency modulation transmission of multiplex telephone signals.

To simplify the treatment, we will assume that the insertion loss of the mode filter is sufficiently large, that is,

$$A_M \rightarrow \infty. \qquad (136)$$

Then, using equations (66), (67), (81), (82), and (83), we obtain

$$d_1 = \frac{\sigma_K{}^4 \cdot N}{l\ \Delta\tau} R_1, \quad d_2 = \frac{\sigma_K{}^4 \cdot N}{l\ \Delta\tau} R_2 \qquad (137)$$

(for spurious modes having no degeneration)

and

$$d_1 = \frac{\sigma_{K,d}{}^4 \cdot N}{2l|\Delta\tau|} R_1, \quad d_2 = \frac{\sigma_{K,d}{}^4 \cdot N}{2l|\Delta\tau|} R_2 \qquad (138)$$

(for spurious modes having even-odd degeneration)

where

$$\begin{aligned} R_1 = \Big[& Q_1 E_4\left(2\left|\frac{\Delta\alpha}{\Delta\tau}\right| t_c\right) \\ & -\frac{5}{6}\left|\frac{t_c}{\Delta t_M}\right| Q_1 E_5\left(2\left|\frac{\Delta\alpha}{\Delta\tau}\right| t_c\right) \\ & + Q_2 E_6\left(2\left|\frac{\Delta\alpha}{\Delta\tau}\right| t_c\right) \\ & -\frac{7}{8}\left|\frac{t_c}{\Delta t_M}\right| Q_2 E_7\left(2\left|\frac{\Delta\alpha}{\Delta\tau}\right| t_c\right) \end{aligned} \qquad (139)$$

$$R_2 = K^2\Big[|\Delta t_M| E_0\left(2\left|\frac{\Delta\alpha}{\Delta\tau}\right| |\Delta t_M|\right)$$

$$- t_c E_0\left(2\left|\frac{\Delta\alpha}{\Delta\tau}\right|t_c\right) - \frac{|\Delta t_M|}{2}E_1\left(2\left|\frac{\Delta\alpha}{\Delta\tau}\right||\Delta t_M|\right)$$

$$+ \frac{t_c^2}{2|\Delta t_M|}E_1\left(2\left|\frac{\Delta\alpha}{\Delta\tau}\right|t_c\right)\Bigg]. \tag{140}$$

For

$$0 < t_c \leq |\Delta t_M|. \tag{141}$$

And

$$R_1 = \Bigg[\widetilde{Q}_1 E_4\left(2\left|\frac{\Delta\alpha}{\Delta\tau}\right||\Delta t_M|\right)$$

$$- \frac{5}{6}\widetilde{Q}_1 E_5\left(2\ \frac{\Delta\alpha}{\Delta\tau}\right||\Delta t_M|\Big)$$

$$+ \widetilde{Q}_2 E_6\left(2\left|\frac{\Delta\alpha}{\Delta\tau}\right||\Delta t_M|\right)$$

$$- \frac{7}{8}\widetilde{Q}_2 E_7\left(2\left|\frac{\Delta\alpha}{\Delta\tau}\right||\Delta t_M|\right) \tag{142}$$

$$R_2 = 0. \tag{143}$$

For

$$0 < |\Delta t_M| \leq t_c \tag{144}$$

$$Q_1 = \frac{2}{5}\pi^4\sigma^2 f_h^2 t_c^5, \quad Q_2 = \frac{8}{21}\pi^6\sigma^4 f_h^2 t_c^7 \tag{145}$$

$$\widetilde{Q}_1 = \frac{2}{5}\pi^4\sigma^2 f_h^2|\Delta t_M|^5, \ \widetilde{Q}_2 = \frac{8}{21}\pi^6\sigma^4 f_h^2|\Delta t_M|^7 \tag{146}$$

$$\Delta t_M = \Delta\tau l_M \tag{147}$$

where l_M is the spacing of the mode filters. When there are no mode filters, we set

$$|\Delta t_M| \to \infty$$

and obtain:

$$R_1[\infty] = \lim_{|\Delta t_M|\to\infty} R_1 = Q_1 E_4\left(2\left|\frac{\Delta\alpha}{\Delta\tau}\right|t_c\right)$$

$$+ Q_2 E_6\left(2\left|\frac{\Delta\alpha}{\Delta\tau}\right|t_c\right) \tag{148}$$

$$R_2[\infty] = \lim_{|\Delta t_M|\to\infty} R_2 = K^2\left(\frac{1}{2}\left|\frac{\Delta\tau}{\Delta\alpha}\right| - t_c E_0\left(2\ \frac{\Delta\alpha}{\Delta\tau}\ t_c\right)\right)$$

$$= \frac{K^2}{2}\left|\frac{\Delta\tau}{\Delta\alpha}\right|e^{-2\left|\frac{\Delta\alpha}{\Delta\tau}\right|t_c}. \tag{149}$$

Naturally equations (148) and (149) agree with the results in section 3.1.

When $\Delta\alpha$ is small and $\hat{\beta}/f_h$ is not very large, we can determine l_M satisfying

$$2\left|\frac{\Delta\alpha}{\Delta\tau}\right||\Delta t_M| = 2\Delta\alpha l_M \ll 1 \tag{150}$$

$$|\Delta t_M| = |\Delta\tau\, l_M \leq t_c. \tag{151}$$

Then equation (142) is approximated with

$$R_1 = \frac{1}{6}\widetilde{Q}_1 + \frac{1}{8}\widetilde{Q}_2. \tag{152}$$

Neglecting $\widetilde{Q}_2$, and by means of equation (146), we get a roughly approximate formula:

$$R_1 \fallingdotseq 0.515\hat{\beta}^2 f_h^2|\Delta t_M|^5, \quad 0 < |\Delta t_M| \leq t_c. \tag{153}$$

For large $\hat{\beta}/f_h$, equations (139) and (149) are approximated with

$$R_1 = Q_1 + Q_2 \fallingdotseq 8.17\frac{1}{f_h}\left(\frac{\hat{\beta}}{f_h}\right)^{-4.25} \tag{154}$$

$$R_2 = \frac{K^2}{2}|\Delta t_M| \fallingdotseq 16.8\left(\frac{\hat{\beta}}{f_h}\right)^{-3}|\Delta t_M|. \tag{155}$$

Neglecting R_1 we obtain:

$$R = R_1 + R_2 \fallingdotseq 16.8\left(\frac{\hat{\beta}}{f_h}\right)^{-3}|\Delta t_M|. \tag{156}$$

We will use equation (153) to discuss the characters of waveguide lines with mode filters, because the insertion of mode filters has a large effect only when $\Delta\alpha$ is small and $\hat{\beta}/f_h$ is not very large.

We will define a suppresion ratio H in order to compare the echo distoriion before and after the insersion of mode filters.

$$H = \frac{R_1}{R_2[\infty]} = 1.03\Delta\alpha\Delta\tau^4\frac{\hat{\beta}^2 f_h^2}{K^2}l_M^5 \tag{157}$$

where we have assumed that $\Delta\alpha$ is so small that we can consider

$$e^{-2\left|\frac{\Delta\alpha}{\Delta\tau}\right|t_c}=1.$$

Equation (157) is valid for $0<|\Delta t_M|\leq t_c$. We will define a characteristic spacing l_c which is the largest l_M satisfying the above restriction. That is,

$$l_c=\frac{t_c}{|\Delta\tau|}. \tag{158}$$

Table 4

H AND l_c FOR MAIN SPURIOUS MODES IN COPPER-PIPE WAVEGUIDES
$N_c=2000$, $f_h=8$ Mc, $\hat{\beta}=29.9$ Mc,
$f_0=50$ Gc, $2a=51$ mm

Mode	$H/l_M{}^5$	$l_c=t_c/\lvert\Delta\tau\rvert$
TE$_{12}$	5.2×10^{-2} (km^{-5})	0.74 km
TE$_{13}$	44	0.16
TE$_{22}$	3.1	0.34
TE$_{23}$	3.8×10^2	0.105
TE$_{02}$	1.8	0.294
TE$_{03}$	3.7×10^2	0.100
TF$_{04}$	11.9×10^3	0.050
TE$_{05}$	2.24×10^5	0.027

In Table 4, $H/l_M{}^5$ and l_c are calculated for the eight main modes in a circular waveguide with diameter of 51 mm when the carrier frequency is 50 Gc, f_h is 8 Mc, and $\hat{\beta}$ is 29.9 Mc.

It can be observed from Table 4 that the insertion of mode filters has a large effect on the low order modes with moderate insertion spacing, but has little effect on the high order modes unless the insertion spacing is very small.

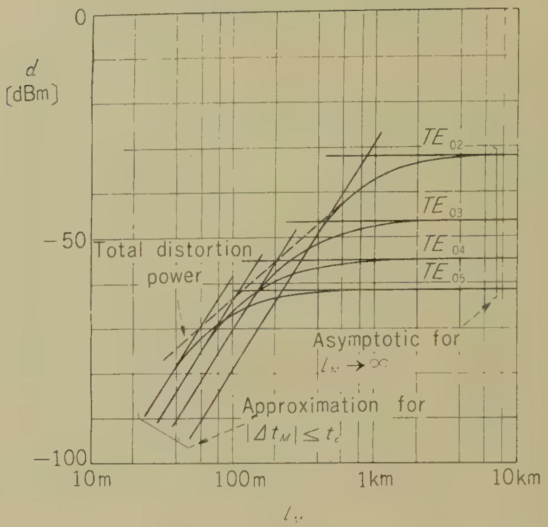


Fig. 14—Relation between echo distortion and spacing of mode filters in helix waveguides. Mode filters are assumed to absorb power of TE$_{02}$, TE$_{03}$, TE$_{04}$, ⋯ ⋯ completely. $N_c=2000$, $f_h=8$ Mc, $\hat{\beta}=29.9$ Mc, $2a=51$ mm, $f_0=50$ Gc, $l=5$ m, $L=20$ km. r.m.s. radius difference at junction: 0.1 mm.

In Fig. 14, the relation between the distortion and the spacing of the mode filters is calculated for the TE$_{0n}$ group of modes in the helix waveguide line. Calculations have been made approximately with equations (149) and (153).

From Fig. 14 we can derive the relation between the waveguide tolerance and the spacing of mode filters in the helix waveguide line. The spurious modes other than TE$_{0n}$ are not taken into account here. The results are shown in Fig. 15.

Though the allowable standard deviation of the waveguide diameter increases as the spacing of mode filters decreases, it is not practical to use very small spacing, because signal mode loss and signal mode reflection due to mode filters can not be ignored when a large number of mode filters is used.

If we take the spacing of 500 m for a practical example, the allowable standard deviation of waveguide diameter becomes about 25 μ.

Table 5

ILLUSTRATIVE EXAMPLES

waveguide diameter: 51mm, length of unit waveguide: 5 m, carrier frequency: 50 Gc

Waveguide	Peak frequency deviation	Occupied bandwidth	Spurious mode to be considered	Waveguide tolerance
Copper-pipe	30 Mc	228 Mc	TE_{12}	r.m.s. junction tilt: 0.17′ r.m.s. junction offset: 5.0 μ
Copper-pipe	75 Mc	500 Mc	TE_{12}	r.m.s. junction tilt: 0.3′ r.m.s. junction offset: 8.5 μ
Helix	30 Mc	228 Mc	TE_{02}	Standard deviation of diameter: 13 μ
Helix	75 Mc	500 Mc	TE_{02}	Standard deviation of diameter: 22 μ
Helix, with TE_{0n} mode filters of 500 m spacing	30 Mc	228 Mc	TE_{02}	Standard deviation of diameter: 25 μ

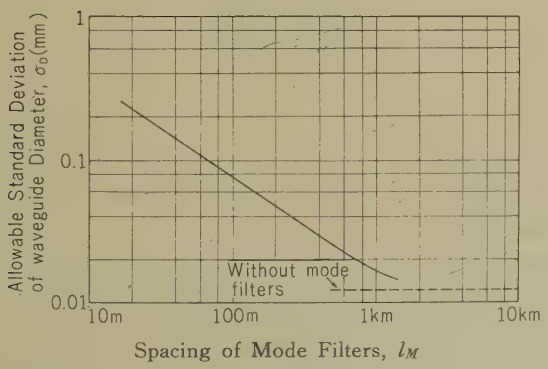


Fig. 15—Waveguide tolerance of helix waveguide line with TE_{0n} mode filters. $N_c=2000$, $f_h=8$ Mc, $\hat{\beta}=29.9$ Mc, $2a=51$ mm, $f_0=50$ Gc.

4. Conclusion

The characteristics of interference waves produced by mode conversion and reconversion in waveguide communication lines have been discussed, and the echo distortion caused by the interference waves in the frequency modulation transmission of multiplex telephone signals has been calculated theoretically. Also, the allowable tolerances of waveguide geometry for frequency modulation communication have been estimated.

Typical examples from the calculations in this paper are illustrated in Table 5.

It is thought that the use of copper-pipe waveguides for frequency modulation communication is very difficult, because highly precise waveguides and waveguide junctions are required. Helix waveguides are more promising, and fairly large diameter tolerance is allowable if large frequency deviation is used. But it is considered from the point of signal transmission that large diameter tolerance cannot be allowed because the increase of signal mode attenuation due to mode conversion becomes very large if the waveguides have poor construction.

The effect of mode filters on the echo distortion has been discussed. The insertion of mode filters which absorb the power of spurious modes other than TE_{0n} was not discussed in detail, because it is considered practical to use helix waveguides which are superior to copper-pipe waveguides with these mode filters.

An example of helix waveguide line with mode filters which could absorb the power of the TE_{0n} spurious modes was calculated, and it was observed that large waveguide tolerance is allowed in this example. But this case is also not practical because we could not realize the concrete construction of these mode filters up to date.

There are many problems to be solved in millimeter wave communication research. For example, even in frequency modulation communication, the research of the distortion due to waveguide bends and mode excitors, the realization of delay equalizers, and the experimental verification of the theoretical investigations should be made in the future.

5. Acknowledgements

The author wishes to thank Dr. I. Someya, Mr. T. Miwa, Dr. B. Oguchi, Mr, T. Fujii, and Mr. F. Ishihara for many helpful discussions and suggestions, and Mr. S. Uchino and Mr. Y. Mori for many suggestions on the application of frequency modulation theory.

Appendix 6.1.
Mode Conversion Coefficients at the Function of the Circular Waveguides

a: radius of circular waveguide
λ: wavelength in free space
λ_g: wavelength in the waveguide
χ: eigenvalue of the mode
$[i, j]$: indication of $TE_{i,j}$ mode
(i, j): indication of $TM_{i,j}$ mode
${}_{T}S^{[0,1]}_{[i,j]}$: transmitting coupling coefficient from TE_{01} mode to $TE_{i,j}$ mode
${}_{R}S^{[0,1]}_{[i,j]}$: reflecting coupling coefficient from TE_{01} mode to $TE_{i,j}$ mode
${}_{T}S^{[0,1]}_{(i,j)}$: transmitting coupling coefficient from TE_{01} mode to $TM_{i,j}$ mode
${}_{R}S^{[0,1]}_{(i,j)}$: reflecting coupling coefficient from TE_{01} mode to $TM_{i,j}$ mode

Small tilt, tilt angle of δ radian, $|\delta|\ll 1$

$${}_{\substack{T\\R}}S^{[0,1]}_{[1,n]} = -jB_{[1,n]}\frac{a(\lambda_{g[0,1]}\pm\lambda_{g[1,n]})^2}{(\lambda_{g[0,1]}\cdot\lambda_{g[1,n]})^{3/2}}\delta$$

$$B_{[1,n]} = \sqrt{2}\,\pi\chi_{[0,1]}\frac{\chi_{[1,n]}}{\sqrt{\chi^2_{[1,n]}-1}}\,\frac{\chi_{[1,n]}}{(\chi^2_{[0,1]}-\chi^2_{[1,n]})^2}$$

Table 6 (a)

CASE OF SMALL TILT

$B_{[1,n]}$

Mode	TE_{11}	TE_{12}	TE_{13}	TE_{14}	TE_{15}	TE_{16}	TE_{17}	TE_{18}
$B_{[1,n]}$	0.295	0.489	0.0432	0.0134	0.00596	0.00320	0.00192	0.00125

Table 6 (b)

CASE OF SMALL TILT

$j{}_{\substack{T\\R}}S^{[0,1]}_{[1,n]}$, $j{}_{T}S^{[0,1]}_{(1,1)}$, $2a=51$ mm, $\lambda=6$ mm, tilt of 1 degree

Mode	TE_{11}	TE_{12}	TE_{13}	TE_{14}	TE_{15}	TE_{16}	TE_{17}	TE_{18}	TM_{11}
Transmitting coupling coefficient	0.0863	0.143	0.0124	0.00375	0.00162	0.000828	0.000470	0.000285	−0.0859
reflecting coupling coefficient	1.4×10^{-6}	3.5×10^{-6}	5.9×10^{-6}	8.7×10^{-6}	1.2×10^{-5}	1.8×10^{-5}	2.7×10^{-5}	4.8×10^{-5}	0

$$ {}_{T}S^{[0,1]}_{(1,1)} = j(1.160)\frac{a}{\lambda}\delta $$

$$ n = 1, 2, 3, \cdots\cdots. $$

Small offset, offset of d, $|d| \ll 1$

$$ {}_{\substack{T\\R}}S^{[0,1]}_{[1,n]} = A_{[1,n]}\sqrt{\frac{\lambda_{g[0,1]}}{\lambda_{g[1,n]}}}\left(\frac{\lambda_{g[1,n]}}{\lambda_{g[0,1]}} \pm 1\right)\frac{d}{a} $$

$$ A_{[1,n]} = -\frac{\chi_{[0,1]}}{\sqrt{2}} \frac{\chi_{[1,n]}}{\sqrt{\chi^2_{[1,n]}-1}} \frac{\chi_{[1,n]}}{\chi^2_{[0,1]}-\chi^2_{[1,n]}} $$

$$ n = 1, 2, 3, \cdots\cdots. $$

Small radius difference, radius difference δ, $|\delta| \ll a$

$$ {}_{\substack{T\\R}}S^{[0,1]}_{[0,n]} = \mp C_{[0,n]}\frac{\lambda_{g[0,1]} \pm \lambda_{g[0,n]}}{2\sqrt{\lambda_{g[0,n]} \cdot \lambda_{g[0,1]}}} \frac{\delta}{a} $$

$$ C_{[0,n]} = \frac{2\chi_{[0,1]} \cdot \chi_{[0,n]}}{\chi^2_{[0,n]} - \chi^2_{[0,1]}} $$

$$ {}_{R}S^{[0,1]}_{[0,1]} = -\frac{1}{2}\left(\frac{\chi_{[0,1]}}{2\pi}\right)^2 \frac{\lambda_g^2{}_{[0,1]}}{a^2} \frac{\delta}{a} $$

$$ n = 2, 3, 4, \cdots\cdots. $$

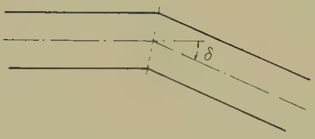


Fig. 16—Tilt angle of δ radian.

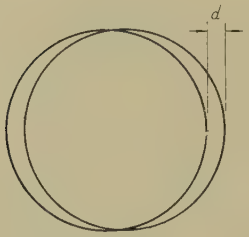


Fig. 17—Offset of d.

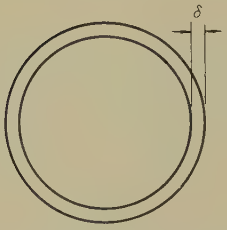


Fig. 18—Radius difference α.

Table 7 (a)

CASE OF SMALL OFFSET

$A_{[1,n]}$

Mode	TE_{11}	TE_{12}	TE_{13}	TE_{14}	TE_{15}	TE_{16}	TE_{17}	TE_{18}
$A_{[1,n]}$	−0.526	1.070	0.400	0.260	0.196	0.158	0.133	0.114

Table 7 (b)

CASE OF SMALL OFFSET

${}_{\substack{T\\R}}S^{[0,1]}_{[1,n]}$, $2a = 51$ mm, $\lambda = 6$ mm, ofiset of 1 mm

Mode	TE_{11}	TE_{12}	TE_{13}	TE_{14}	TE_{15}	TE_{16}	TE_{17}	TE_{18}
Transmitting coupling coefficient	−0.0413	0.0839	0.0314	0.0204	0.0154	0.0125	0.0107	0.00987
Reflecting coupling coefficient	0.00017	0.00042	0.00068	0.00098	0.0014	0.0018	0.0025	0.0040

Table 8 (a)

CASE OF SMALL RADIUS DIFFERENCE

$C_{[0,n]}$

Mode	TE_{02}	TE_{03}	TE_{04}	TE_{05}	TE_{06}	TF_{07}	TF_{08}
$C_{[0,n]}$	1. 557	0. 878	0. 627	0. 492	0. 406	0. 347	0. 302

Table 8 (b)

CASE OF SMALL RADIUS DIFFERENCE

$-{}_{\substack{T\\R}}S^{[0,1]}_{[0,n]}$, $2a=51$ mm, $\lambda=6$ mm, $\delta=1$ mm

Mode	TE_{02}	TE_{03}	TE_{04}	TE_{05}	TE_{06}	TE_{07}	TE_{08}	TE_{01}
Transmitting coupling coefficient	0. 0611	0. 0345	0. 0246	0. 0194	0. 0162	0. 0143	0. 0133	
Reflecting coupling coefficient	0. 000412	0. 00072	0. 0012	0. 0016	0. 0022	0. 0030	0. 0044	0. 000412

Table 9 (a)

CASE OF SMALL ELLIPTICAL DEFORMATION

$D_{[2,n]}$

Mode	TE_{21}	TE_{22}	TE_{23}	TE_{24}	TE_{25}	TE_{26}	TE_{27}	TE_{28}
$D_{[2,n]}$	1. 448	−0. 443	−0. 230	−0. 161	−0. 125	−0. 100	−0. 087	−0. 076

Table 9 (b)

CASE OF SMALL ELLIPTICAL DEFORMATION

${}_{\substack{T\\R}}S^{[0,1]}_{[2,n]}$, $2a=51$ mm, $\lambda=6$ mm, $\Delta=1$ mm

Mode	TE_{21}	TE_{22}	TE_{23}	TE_{24}	TE_{25}	TE_{26}	TE_{27}	TE_{28}
Transmitting coupling coefficient	0. 161	−0. 0482	−0. 0255	−0. 0179	−0. 0139	−0. 011	−0. 010	−0. 011
Reflecting coupling coefficient	−0. 00032	−0. 00054	−0. 00086	−0. 0011	−0. 0015	−0. 0021	−0. 0031	−0. 0062

Small elliptical deformation,

$$
{}_{T}\overset{[0,1]}{\underset{R\ [2,n]}{S}} = D_{[2,n]}\frac{\lambda_{g[2,n]}\pm\lambda_{g[0,1]}}{\sqrt{\lambda_{g[0,1]}\cdot\lambda_{g[2,n]}}}\sqrt{2}\,\frac{\Delta}{a}
$$

$$
D_{[2,n]} = \frac{\chi_{[0,1]}\cdot\chi^2{}_{[2,n]}}{2\sqrt{\chi^2{}_{[2,n]}-4\left(\chi^2{}_{[0,1]}-\chi^2{}_{[2,n]}\right)}}.
$$

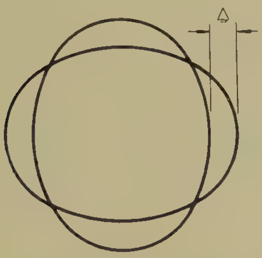


Fig. 19—Electrical deformation Δ.

Appendix 6.2.
Calculation of $C(N, n, m)$

$C(N, n, m)$ is the number of cases when we select n selection from a sequence of N sections connected in cascade, and there are edges of these n selected sections, provided we take no account of the point as the edge where the two selected sections adjoin each other.

We will set

$$m = 2r$$

because m is an even integer.

We will consider a sequence of n sections connected in cascade, and will divide it into r parts. The number of these cases is:

$${}_{n-1}C_{r-1}.$$

These r parts are separated and the remaining $N-n$ sections are distributed between and sides of them. The number of these cases is:

$${}_{N-n+1}C_{r}.$$

Then we obtain:

$$C(N, n, m) = {}_{n-1}C_{\frac{m}{2}-1}\cdot{}_{N-n+1}C_{\frac{m}{2}}.$$

Appendix 6.3.
Characteristics of Function $E_n(a)$

Definition:

$$E_n(a) = (n+1)\int_0^1 u^n e^{-au}du. \tag{1}$$

Here we will assume:

$$n = 0, 1, 2, 3, \cdots\cdots$$

$$a \geq 0.$$

Then the following equations are easily derived:

$$E_n(a) > 0 \quad \text{for} \quad \infty > a \geq 0 \tag{2}$$

$$\frac{dE_n(a)}{\partial a} = -\frac{n+1}{n+2}E_{n+1}(a) \tag{3}$$

$$E_n(0) = 1. \tag{4}$$

Therefore, $E_n(a)$ is 1 at $a=0$, and decreases monotonically with increasing a.

Applying Taylor series expansion. we obtain:

$$\int_0^1 u^n e^{-au}du = \frac{1}{n+1} - \frac{a}{n+2} + \frac{a^2}{2(n+3)} - \frac{a^3}{3!(n+4)} + \cdots\cdots.$$

Then we obtain:

$$E_n(a) = 1-\left(\frac{n+1}{n+2}\right)a+\frac{1}{2!}\left(\frac{n+1}{n+3}\right)a^2 - \frac{1}{3!}\left(\frac{n+1}{n+4}\right)a^3+\cdots\cdots. \tag{5}$$

Equation (5) is useful for numerical calculation for small a.

Applying the partial integration formula, we obtain:

$$\int_0^1 u^n e^{-au}du = -\frac{e^{-a}}{a} + \frac{n}{a}\int_0^1 u^{n-1}e^{-au}du$$

$$n = 1, 2, 3\cdots\cdots.$$

Then we obtain following recurrent formulas:

$$\frac{a}{n+1}E_n(a) = E_{n-1}(a) - e^{-a} \quad n = 1, 2, 3, \cdots\cdots \tag{6}$$

and

$$\int_0^1 e^{-au}du = \frac{1-e^{-a}}{a}$$

$$E_0(a) = \frac{1-e^{-a}}{a} \tag{7}$$

for $n=0$.

Applying equation (6) repeatedly, we obtain:

$$E_n(a) = \frac{(n+1)!}{a^{n+1}} - e^{-a}\left\{\frac{n+1}{a} + \frac{(n+1)n}{a^2} + \frac{(n+1)n(n-1)}{a^3} + \cdots\cdots + \frac{(n+1)n(n-1)\cdots\cdots3\cdot2\cdot1}{a^{n+1}}\right\} \tag{8}$$

or

$$E_n(a) = \frac{(n+1)}{a^{n+1}}\left\{1 - e^{-a}\left\{1 + a + \frac{a^2}{2!} + \cdots\cdots + \frac{a^n}{n!}\right\}\right\}. \tag{9}$$

If a is sufficiently larger than n, we can consider:

$$\left(1 + a + \frac{a^2}{2!} + \cdots\cdots + \frac{a^2}{n!}\right)e^{-a} \ll 1.$$

Thus, we obtain the following asymptotic formula:

$$E_n(a) \sim \frac{(n+1)!}{a^{n+1}}, \quad \text{for} \quad n \ll a. \tag{10}$$

A function $f(u)$, defined by

$$\begin{aligned} f(u) &= 0, & u < 0 \\ &= (n+1)u^n, & 0 \leq u \leq 1 \\ &= 0, & u > 1 \end{aligned}$$

behaves like a delta-function at $u=1$, if n is sufficiently large. Therefore we obtain:

$$E_n(a) \sim e^{-a}, \quad \text{for} \quad a \ll n. \tag{11}$$

$E_n(a)$ is calculated numerically for $0.1 \leq a \leq 100$ and $n=0, 1, 2, \cdots\cdot, 8$ in Fig. 20.

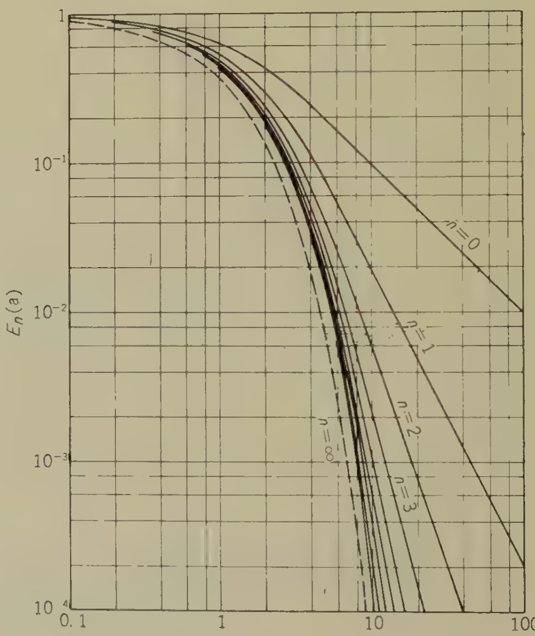


Fig. 20—Function $E_n(a)$.

References

(1) Larsen, H: Delay Distortion and Equalization in H_{01} Waveguide for Long-Range Communication, Convention on Long-Distance Transmission by Waveguide. London (1959.01)

(2) Albersheim W. J. and Shafer J. P.: Echo Distortion in the FM Transmission of Frequency-Division Multiplex. *Proc. I.R.E.*, **40**(3), March 1952, p. 316–328.

(3) Medhurst R. G. and Small G. F.: An extended Analysis of Echo Distortion in the FM Transmission of Frequency-Division Multiplex. *P.I.E.E.*; **103**(8), pt. B, March 1956.
(4) Iiguchi, S.: Mode Conversion in the Transmission of TE_{01} Wave Through a Slight Tilt and a Slight Offset of Waveguide, International Conference on Ultra-High Frequency Circuits and Antennas, Paris (1957.10).
(5) Sugawara and others: FM Radio Engineering, Industrial Daily News Co. Tokyo (1959. 10).

* * * *

Review of the Electrical Communication Laboratory
Volume 9, Numbers 5-6, May-June, 1961

U.D.C. 621.395.344.6

The C8 Crossbar Toll Switching System*

Toshiichi MASUDA†

This paper describes process of the development of toll switching system and outline of The C8 Crossbar Toll System. This System has been developed for the purpose of realization of Nationwide Customer Toll Dialing Service. It has resister function of dial code information and it is a common control system of a maker type with four-wire talking path.

Introduction

In order to realize the nationwide subscriber dialing plan, it is necessary to establish an automatic charging system for long-distance non-delay calls, and thus to promote the expansion and automation of intertoll trunk networks economically.

The C8 system descrbed in this paper is an automatic toll switching system that has been developed for the above purposes. As a result of its invention, an effective way to mechanize transit switching has become feasible.

The C8 system is a common-control system of the marker type with four-wire talking paths. The design of the system was started in 1953, and was completed in 1959 with the cooperation of the Nippon Telegraph and Telephone Public Corporation and the Nippon Electric Company, Ltd.

The first installation was put into service at the Sendai Toll Office in November, 1959. Presently, larger scale installations are being instllael first in Osaka and then in Tokyo Toll Offices. Within several years from now, the C8 system will be adopted at almost all the main cities in Japan and will be the basis of the nationwide dialing networks.

1. The Crossbar Toll System and Intertoll Trunk Networks

2. The Crossbar Toll System and Intertoll Trunk Networks

The basic principles for the future construction of telephone circuits are based on the following goals of service:

(1) any of two subscribers throughout the country may be connected with one another by means of the subscriber toll dialing service

(2) The selection dial code in the subscriber toll dialing service should be based on the simplest Universal Numbering System.

In order to realize the above goals, it is necessary to establish the following plans:

(a) zone system plan
(b) nation-wide numbering plan
(c) transmission loss distrbution plan
(d) charging plan.

From these plans, the functions which the switching system must perform will be evaluated.

According to the transmission loss distribution plan, the station-to-station transmission performance rating is set as 49 dB to 90 per cent of the subsribers. Thus, the net loss of the intertoll trunks which are situated higher than the District Center are to be 0 dB. And at the Regional and District Centers, the four-wire connection is introduced.

In line with the other plans, many toll offices perform the automatic alternate routng function and the dial code register function. The destination code base, the code of which

---

* MS received by the Electrical Communication Laboratory on 1 Aug., 1960. Originally published in the *Kenkyu Zituyôka Hôkoku* (*Electrical Communication Laboratory Technical Journal*), N.T.T., Vol. 10, No. 4(II), pp. 537–586, 1960.

† Switching Section.

does not have to be varied according to the selected route, will be used as the selection code between toll offices.

The present SXS system used alone cannot satisfy the above conditions because of the difficulty of the construction of networks.

Therefore, the crossbar system has been introduced as a toll switching system.

The main characteristics and merits of the crossbar toll switching system are described below.

i) Since alternate routing becomes possible and trunk efficiency can be increased, the network can economically constructed.

ii) The numbering plan becomes simpler and the network can be constructed independent of the numbering plan.

iii) There is very little noise because precious

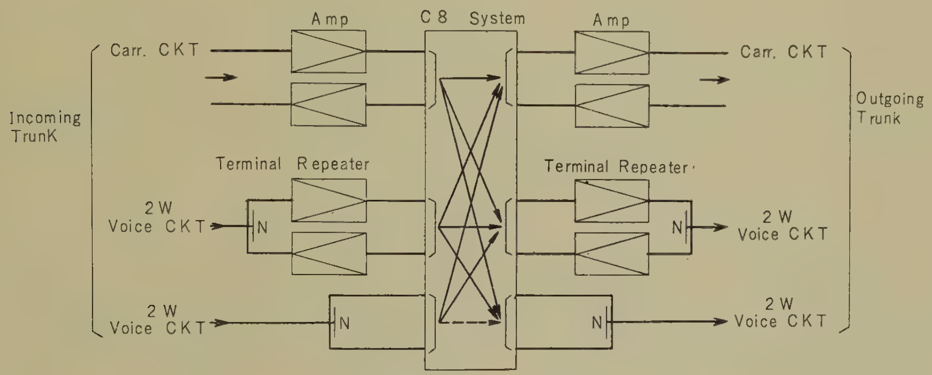


Fig. 1—The 4W. CKT. and 2W. CKT. in the C8 system.

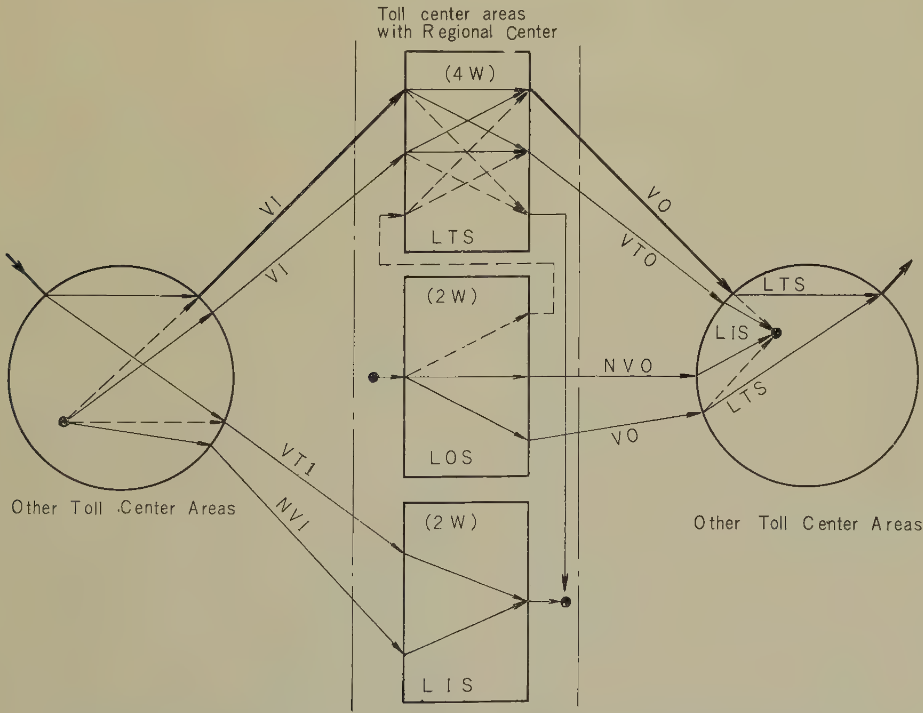


Fig. 2—Intertoll trunks coming into and out of the crossbar toll system.

Table 1

FRAME TYPE AND CAPACITY

Frame Type		Capacity			Max. Trk. Cap.		Note
		Frame Number	Traffic (HCS)	Fundamental Terminal	Incoming	Outgoing	
10 Frame Type	1 Train	2~10	6000~ 30000	200~1000	3000	3000	for small offices
20 Frame Type	1 Train	4~20	12000~ 60000	400~2000	6000	6000	for medium offices
	2 Train	8~40	24000~120000	800~4000	6000	12000	
40 Frame Type	1 Train	8~40	24000~120000	800~4000	12000	12000	for large offices
	2 Train	16~40	48000~240000	1600~8000	12000	24000	

metal contacts are used in the transmission path. Also, four-wire speech paths can be easily constructed so that the transmission design of intertoll trunks becomes easier.

For the above reason, toll switching system such as the C8, C5, and C6 have been developed. The C8 system is mainly used as a transit switch, whereas the C5 and C6 systems, with two wire speach paths, are employed as the incoming and outgoing switches. The C5 and C6 systems have the same functions. but the C6 system is more suitable than the C5 for offices with large capacities.

In these crossbar toll switching systems, the outgoing route is in general identified by the first three or four digits of the received dial codes. These systems can accomodate a maximum of several hundred outgoing routes. Since the selection of the outgoing trunk group can be performed on the full-availability basis the number of trunks needed in the office is fewer than that required in the SXS and other system.

In order to economize the cost of the trunk equipment in the network construction, high usage groups and final groups are established, and the calls which overflow the high usage group are made to select the final group automatically.

In these crossbar toll switching systems, the alternate routing generally resorts to the Far-to-Near Rotation Base.

The multiple-frequency dialing system is adopted in the intertoll trunk as far as possible to increase the efficiency, to economize on the number of reister-sender, required, and to stabilize the supervisory signals.

2. Functions of the C8 System

The kinds of switching systems of the party offices to which the C8 system must be connected vary, and the sizes of the trunk groups range widely from large through medium to small ones. Since there may be frequent changes in the signaling condition due to changes in switching systems of the party offices, size of the trunk groups, routes, and transmission systems; the C8 system is constructed with flexible function units.

The C8 system has been designed with reference to the 4A toll switching system of U.S.A. It is a crossbar system of the common control type, with four-wire connection and four stages. Depending on the capacity of the office, the C8 system is constructed as the 10-frame type, 20-frame type, or 40-frame type. The 20-frame type and 40-frame type can be further expanded into the 2-train type, which doubles the office capacity.

The following is an explanation of the main function of the C8 system.

(1)Switching

It can be used as a long haul transit and as a long haul incoming switch in the regional and district centers; and if CAMA system is attached to it, it can be used as a long haul outgoing switch.

(2) Trunk

It can be accomodated either a one way or both way intertoll trunk by means of the toll dialing.

(3)Selection Codes

The selection codes are the total digits used in the nation-wide dialing. The signaling type is either the multi-frequency type or the dial pulse type. The maximum number of digits which can be stored in the register is 8.

The selection of the outgoing trunks is performed by the first three or four digits.

(4) Transmission

The zero net loss connection is preformed by means of the four-wire connection. When a two-wire trunk is used, attachment of a hybrid coil to maks it function as a four-wire trunk at the switching point. The PAD control function can be added whenever necessary.

(5)Alternate routing

At most, the C-8 system is capable of automatic alternate routing over three routes

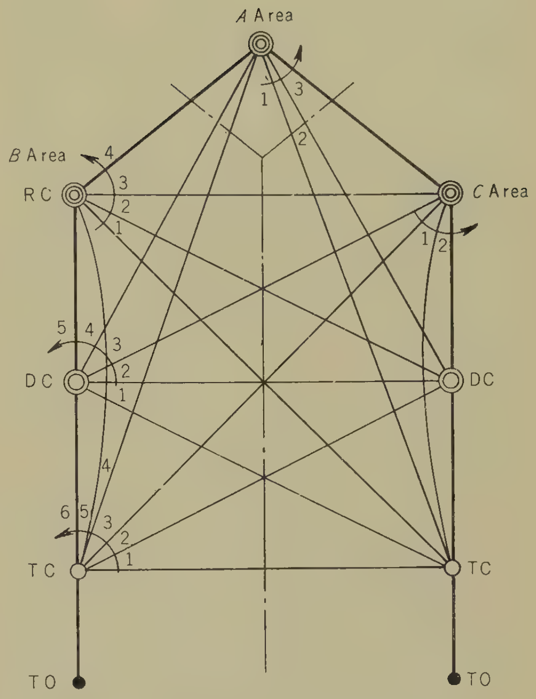


Fig. 3—Alternate routing pattern.

(1) This illustrates how the connection is made between the B area TC and the C area TC.

(2) The arrows designate the direction ot the alternate routing, with the numbers showing its order.

(3) The heavy line stands for the final group and the fine lines represents the high usage groups.

3. The C8 Equipment

The compoition and the connections of the C8 system will be described first.

The transmission path is composed of the crossbar switch, and of incoming and outgoing trunks. On the other hand, the register-sender, the decoder, the translator, and the marker control the establishment of the connection in which automatic alternate routing etc. are carried out.

A call through the incoming trunk, after passing through the sender link, seizes the register-sender, and sends the selection codes into the register-sender to be registered. The register-sender lets the decoder, the translator and the marker start, and sends the initial part of the selection codes into these. The decoder translates the selection codes into operational information by means of the translator, and sends them to the marker. According to this information, the marker selects the outgoing trunk, operates the crossbar switch, and establishes the transmission path. The register-sender sends the selection codes through the crossbar switch to the backward offices. The approximate holding time of the register-sender is 10 seconds, that of the marker and the decoder together is 0.7 seconds, and that of the translator 0.05 seconds.

Incidentally, for the purpose of stabilizing and prolonging the life of the equipment, wire spring relays are adopted throughout the C8 system, in common with other crossbar systems. Also the solderless wrapped connection method is adopted for the same purpose.

Next the outline of the main parts of the equipment is explained.

(1) Switch Frame

The five-wire crossbar switch is used as the main switch. Of the five leads, four are for talking and the one is for holding.

The incoming frame and the outgoing frame are both composed of two stages. This means that the whole set consists of four stages.

The capacity of one frame is 100 trunks in the basic construction, and it is capable of carrying approximately 3000 HCS traffic. When the frame traffic does not reach the above value because of low efficiency in the usage of trunks, trunk extension frames are further attached in order to increase the number of trunks to be accomodated.

(2) Register-sender

There are two types of the register-senders. One is the MF (Multi-frequency) type, the other is the DP (Dial Pulse) type. The former is connected with the MF type incoming trunk, and the latter with the DP type trunk. Both of them can send either the DPor MF codes.

Selection codes with a maximum of 8 digits can he registered. In the future, however,

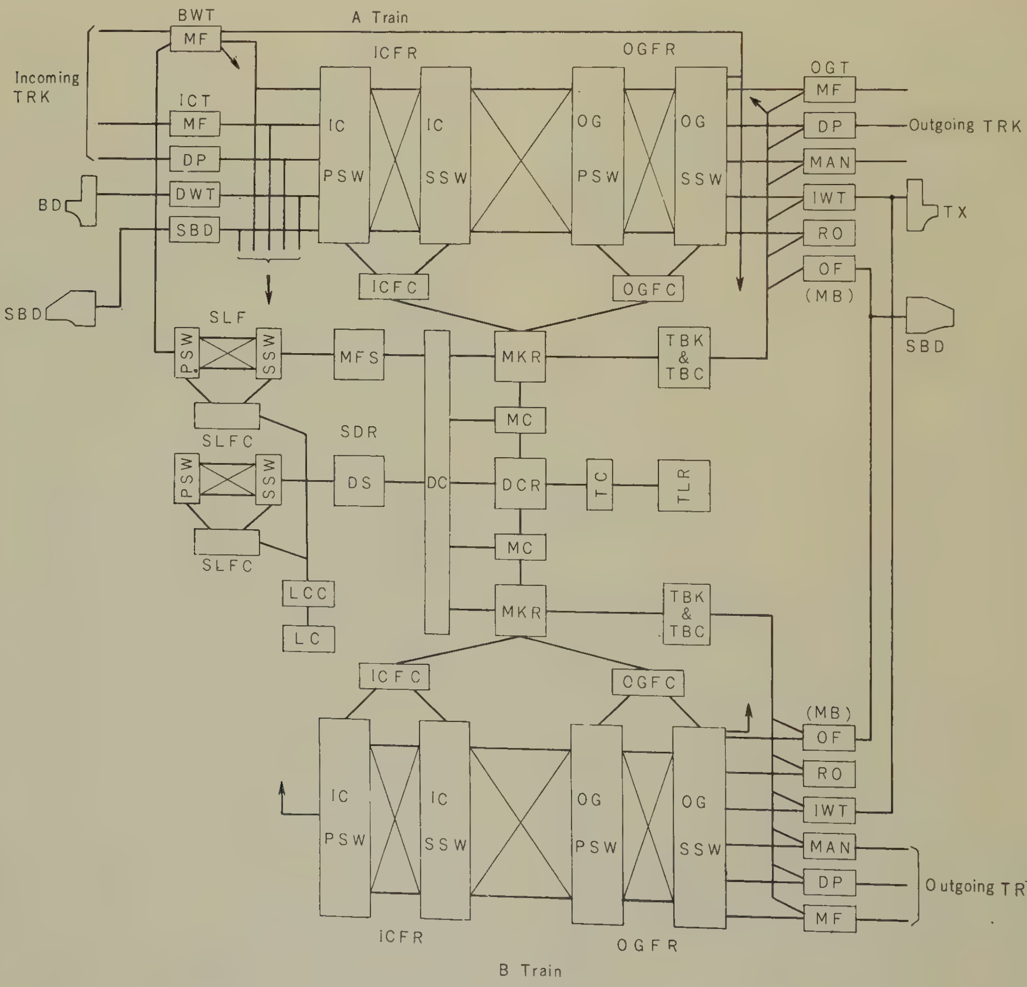


Fig. 4—Fundamental connection scheme (2 train.)

this will he increased to 9 according to the numbering plan.

The register-sender receives and registers at its registering part the codes sent from the forward offices. After the transmission path is completed, the necessary number of digits are sent through it. Furthermore the register-sender pre-translates the first two digits of the codes and decides the starting condition of the decoder.

(3) Decoder and Translator

The decoder and the translator cooperatively perform the automatic alternate routing function.

The translator is a memory equipmet of the ring coil type. It can store about 2000 kinds of information, which are designated by jumper leads.

The decoder receives the information about the outgoing route which should be selected first. The decoder then carries out a group busy test on each trunk group in the first route. Selects the trunk group which contains idle trunks, acquires the information about that trunk group from the translator, and sends the information to the marker. When there is no idle trunk in the first outgoing route, the decoder automatically proceeds to the group busy test on the second routes. In this way the decoder and the translator can execute automatic alternate routing up to three times at most.

(5) Marker

The maker connects the necessry outgoing trunk groups according to the information acquired from the translator, using the trunk block connector. The maker then carries out the busy test and seizes the trunk which should be used. It also identifies the accomodated switch frame number of the incoming and outgoing trunk by virtue of the MF type frame number identifier. The maker then selects and seizes the idle link between the incoming and outgoing trunk operates the crossbar switch, and completes the transmission path

(5) Trouble Recorder

All trouble that are detected in the common control equipment are sent to the trouble recorder and punched into a tape. A sixty unit puncher is used in the trouble recorder. The total information capacity that can be punched in one cycle of operation is about are thousand unit.

Conclusion

In order to realize the nationwide subscriber dialing plan, a new toll switching system, called the C8 crossbar switching system, has been developed. It is a four-wire transit switching system using crossbar switches and the wire spring relays, and will be used at the main switching points of the long-distance non-delay networks in Japan.

This system can also be connected to the electronic CAMA system which is now being devloped. When connected to the CAMA system, it can be used as the long haul outgoing switch in the nationwide subscriber dialing service.

Acknowledgement

The author would like to express a deep appreciation to the members of the Nippon Electric Company, Ltd., who have cooperated with him during the development of the C8 crossbar system.

Reference

(1) "Outline of Crossbar Toll Syitching System," *Electr. Comm. Lab. Tech. Jour.*, Vol. 8 No. 6, 1959.

* * * *

Papers Contributed to Scientific and Technical Journals by the Members of the Laboratory

Papers Published in Other Publications of the Electrical Communication Laboratory

U.D.C. 537.311.4.08 : 546.3 : 539.23

Contact Resistance due to Tunnel Effect and Resistance of Metallic Films
Yoshihiro MIZUSHIMA and *Osamu OCHI*
Kenkyû Zituyôka Hôkoku (Electr. Comm. Labor. Techn. Journ.), NTT, **10**, 3, pp. 353–367, Mar. 1961

Tunnel-resistance between metallic contacts was calculated to show voltage and temperature dependencies. Typically a semiconductive characteristic was found. The resistance of very-thin evaporated films is compared with this contact resistance.

U.D.C. 537.562 : 538.569.029.63

Measurements of Plasma Density by Microwave
Kazuo TAKAYAMA, *Seiji OHARA*, and *Hideo IKEGAMI*
Kenkyû Zituyôka Hôkoku (Electr. Comm. Labor. Techn. Journ.), NTT, **10**, 4(1), pp. 411–428, Apr. 1961

The theoretical and experimental aspects of measuring electron densities of a plasma column in a waveguide by microwave methods are discussed. The results agree with those obtained from the probe method.

U.D.C. 538.561 : 621.318.57 : 621.3.8.1
621.395.301 : 621.374.32.001.2

8-Magnetic-Core Super Parametron
Bunpei OKADA and *Tamio SHIMIZU*
Kenkyû Zituyôka Hôkoku (Electr. Comm. Labor. Techn. Journ.), NTT, **10**, 3, pp. 299–313, Mar. 1961

A study and the results of experiments on controlled oscillations in 8-magnetic-core parametrons are described. Applications in an electronic ZZZ system are also explained.

U.D.C. 621.315.615 : 665.335.5 : 621.319.414

Castor-Oil Impregnated Paper Capacitors
Koichi KUDO
Kenkyû Zituyôka Hôkoku (Electr. Comm. Labor. Techn. Journ.), NTT, **10**, 4(1), pp. 495–510, Apr. 1961

This paper presents the dielectric properties of castor-oil and dehydrated castor-oil. The effects of these properties on impregnated paper capacitors are also discussed.

Translations of many of these papers will appear in a future addition of the Review of Electrical Communication Laboratory.

Please address requests for copies of these papers, which are in Japanese, directly to the authors to whom credited at: The Electrical Communication Laboratory, 1551 Kitizyôzi, Musasino-si, Tokyo.

U.D.C. 621.315.619 : 678.652

Study of Internally Plasticized Melamine Laminates
Masaaki KALAAKI and *Kazuhiro KOYAMA*
Kenkyû Zituyôka Hôkoku (Electr. Comm. Labor. Techn. Journ.), NTT, **10**, 4(1), pp. 511–525, Apa. 1961

The properties of a new melamine laminate which is made from cotton paper and internally plasticized melamine resin are described. This new melamine laminate can be punched and postformed.

U.D.C. 621.317.333.8 : 621.315.212 : 621.317.34

Pulse Echo Meter for Long Distance Coaxial Cables
Yosio YOKOSE and *Ken'ichi HAYASI*
Kenkyû Zituyôka Hôkoku (Electr. Comm. Labor. Techn. Journ.), NTT, **10**, 4(1), pp. 527–535, Apr. 1961

A new pulse echo meter for measuring the impedance irregularities of long distance coaxial cables is described' Transmission distortion of test cables is equalized by equalizing networks.

U.D.C. 621.372.54 : 621.316.726 : 621.396.93

CR Selective Amplifier Applied for Mobile Radio System
Takahiro MORINAGA
Kenkyû Zituyôka Hôkoku (Electr. Comm. Labor. Techn. Journ.), NTT, **10**, 3, pp. 339–352, Mar. 1961

This paper contains a theoretical treatment of selective amplifiers using twin T networks. Experimental measurements on the stability of these amplifiers are also discussed.

U.D.C. 621.372.543.2

Design of Zig-Zag Filter
Noboru KUROSAWA and *Kunio TASIRO*
Kenkyû Zituyôka Hôkoku (Electr. Comm. Labor. Techn. Journ.), NTT, **10**, 4(1), pp. 453–475, Apr. 1961

The theory and design of the Zig-Zag filter, which is used for band pass filters, are described. Diagrams and equations are given, and their practical use is explained.

U.D.C. 621.375.9 : 621.382.2

Construction and Properties of Silver Bonded Diode
Shoichi KITA
Kenkyû Zituyôka Hôkoku (Electr. Comm. Labor. Techn. Journ.), NTT, **10**, 4(1), pp. 477–494, Apr. 1961

The silver bonded diode, which is constructed of n-type germanium and a silver-gallium whisker, is described. It is very suitable for microwave parametric amplifiers, because its barrier capacitance is smaller and its cutoff frequency is higher than that of other diodes.

U.D.C. 621.395.36 : 621.395.345
681.142 : 621.374.32 : 621.318.1

Magnetic-Core Memory Used in Electronic ZZZ System
Bunpei OKADA and *Masaru HASEGAWA*
Kenkyû Zituyôka Hôkoku (Electr. Comm. Labor. Techn. Journ.), NTT, **10**, 3, pp. 285–297, Mar. 1961

A study on a two-frequency magnetic-core memory used in an electronic ZZZ system is described. Experimental results are also given.

U.D.C. 621.395.361 : 621.318.57 : 621.318.1

Electronic ZZZ System
Hideo KAWASAKI, Ken-ichi FUKUI, Kazushi ONOSE, Bunpei OKADA,
Toshikazu HAMANO, and *Masaru HASEGAWA*
Kenkyû Zituyôka Hôkoku (Electr. Comm. Labor. Techn. Journ.), NTT, **10**, 3, pp. 251–283, Mar. 1961

An experimental electronic ZZZ equipment using parametrons and magnetic-core matrix-memory circuits is discussed. Design details and the results of a performance test are described.

U.D.C. 621.396.61 : 621.385.623.5 : 621.371 : 551.510.52

Hyper-Power UHF Klystron Transmitter Used for Transmission Beyond the Horison
Shin'ici NAKAMURA
Kenkyû Zituyôka Hôkoku (Electr. Comm. Labor. Techn. Journ.), NTT, **10**, 1, pp. 315–338, Mar. 1961

Some considerations on the adjustment and performance of a hyper-power uhf klystron transmitter which can be used in a 60～120 channel repeater or T.V. repeater are described.

U.D.C. 681.142 : 518.5 : 519.2

Majority Decision Function of Six Variables
Iwao TODA, Mamoru KONDO, and *Saburo MUROGA*
Kenkyû Zituyôka Hôkoku (Electr. Comm. Labor. Techn. Journ.), NTT, **10**, 3, pp. 369–403, Mar. 1961

A method for obtaining the majority decision functions of n variables is proposed. A table of the functions of six variables is shown in this paper.

U.D.C. 681.613.4.001.2 : 007 : 159.938

A Psychological Study on the Typewriting Behavior of Telegraph Operators
Yoshikazu ŌWAKI and *Kin'ya MARUYAMA*
Kenkyû Zituyôka Hôkoku (Electr. Comm. Labor. Techn. Journ.), NTT, **10**, 4(1), pp. 429–452, Apr. 1961

An attempt to construct a teleprinter with Chinese characters, and analysis of the typewriting speed for finger and hand are described.

* * * *

Vol. 9 Nos. 5-6 MAY-JUNE

CONTENTS

Majority Decision Functions of Six Variables 253
Iwao TODA, Mamoru KONDO, and
Saburo MUROGA

Ferroelectric Polarization Reversal by Linear Fields 279
Hiroyuki SHIBATA, Hiroo TOYODA, and
Shigeru WAKU

Formation of Cesium Antimonide. II. Formation of Cs_3Sb at Temperatures below 130° C 287
Kiyoshi MIYAKE

Formation of Cesium Antimonide. III. Logarithmic Growth of Cesium Antimonide Films on Antimony 295
Kiyoshi MIYAKE

A Method of Quantifying Fe^{2+} and Mn^{3+} in Sintered Mn-Zn Ferrites 300
Hideo NAGATO

An Input Admittance Meter for Electron Tubes 309
Takuya KOJIMA and *Hiroshi HARA*

Characteristics of Fading due to Ducts and Quantitative Estimation of Fading 319
Shigetaka UGAI

Interference Waves in Waveguide Transmission and Intermodulation Noise in FM System for Waveguide Communication 361
Kazuhiro MIYAUCHI

The C8 Crossbar Toll Switching System 394
Toshiichi MASUDA

Papers Contributed to Scientific and Technical Journals by the Members of the Laboratory 400

Copyright 1961 by the Electrical Communication Laboratory, NTT